In: 37th International Symposium on Robotics (ISR 2006) / 4th German Conference on Robotics (Robotik 2006)
München, Germany My 15 to 17, 2006

# Acoustic Detection of Contact State Changes of Deformable Linear Objects

Jan Deiterding
University of Bayreuth
Chair for Applied Computer Science III
(Robotics and Embedded Systems)
Germany

Antoine Schlechter
University of Bayreuth
Chair for Applied Computer Science III
(Robotics and Embedded Systems)
Germany

Dominik Henrich
University of Bayreuth
Chair for Applied Computer Science III
(Robotics and Embedded Systems)
Germany

**Speaker:** Dominik Henrich, University of Bayreuth, Chair for Applied Computer Science III (Robotics and Embedded Systems), Geschwister-Scholl-Platz 3, D-95445 Bayreuth, dominik.henrich@uni-bayreuth.de

**Topic:** Research and Development
**Keywords:** Assembly, Contact States, Sensors

## Introduction

In assembly, objects have to be brought in contact with other objects. In order to deal with the imprecisions in the geometry of the mating objects and in the manipulating system, a lot of research concentrates on describing assembly tasks using contact states and on finding corresponding algorithms to detect contact state transitions or even to determine the current contact state based on sensor data [1-3], [6], [8], [9], [13]. Typical force-based approaches only evaluate the low frequency effects of changes in the contact situation. Most often, contact forces are measured to decide on contact state changes using simple thresholds. The sound caused by collisions or by different friction coefficients in different contact situations is not evaluated. In this paper, the sound and the vibration generated by a contact is used to "hear" changes in contact situations. The sound may be recorded using an ordinary microphone. However, the use of a microphone is impractical as the noise of the whole enviroment is recorded as well. As this noise can overlay the signal of interest, the microphone must be placed very close to the expected point of contact. Because of that we use an impact sound sensor that only records the vibration of an object in touch with the sensor and is insensitive to airbourne sound. This sensor is mounted on the robots gripper so we can measure the impact sound of the held object directly, without having to place one or more microphones in the workspace (Fig. 1, left). Using this setup we will try to detect three typical changes in the contact situation between a leaf spring and its environment.

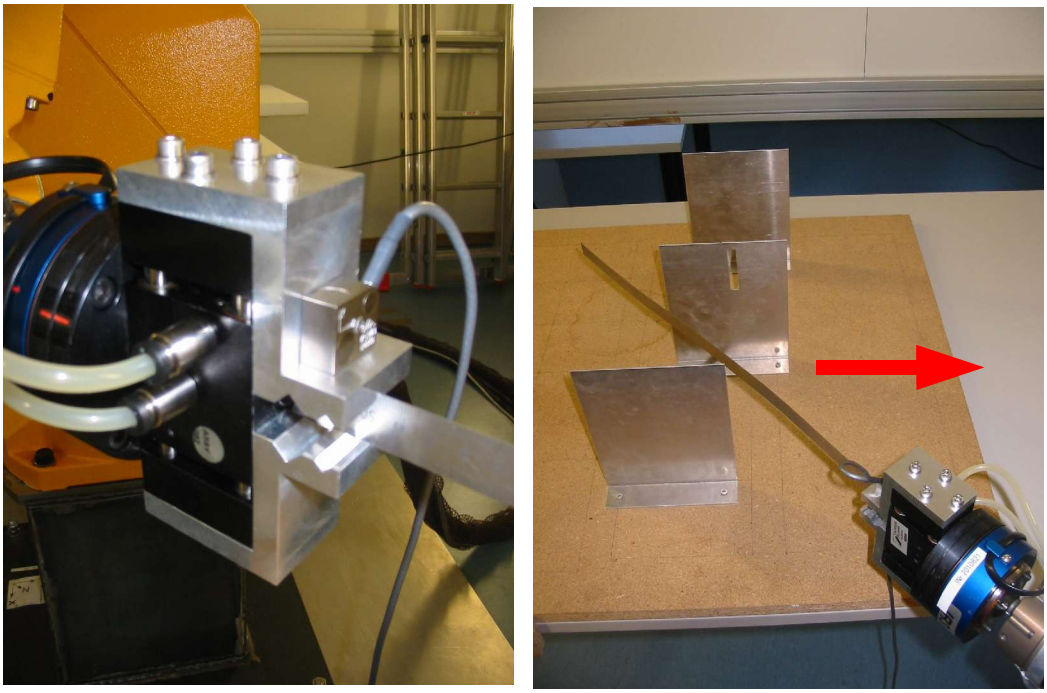

*Fig. 1: Left: Photo of the impact sound sensor mounted on the gripper. Right: Photo of the experimental setup. The endpoint of a leaf spring is moved into contact with the plate, drawn over the slot, and finally off the plate.*

**Experimental setup**

The robot gripper holds a leaf spring that is 50cm long and 2cm wide. The endpoint of the spring is moved into contact with a metal plate, then it is drawn over the plate (Fig. 1, right). In the middle of the plate, the endpoint shortly drops into a slot, with the spring colliding with the inner edge of the slot. Finally the endpoint definitely falls off the plate. The sequence of contact state transitions can be described using the model from [5] that enumerates all possible single contact situations between a deformable linear object (DLO) and a rigid convex polyhedron and analyzed the possible transitions between these contact states [7]. In this model, a DLO is described as an edge E with a free endpoint as one vertex V. Convex polyhedrons consist of faces F, edges E, and vertices V. Single contact states may then be described by the respective elementary parts in contact, e.g., V/F stands for a vertex/face contact. In the following, we always denote the DLO contact part first. If a contact state remains unchanged during a small gripper movement we call it *stable*, otherwise it is called *instable*. *Initiated* contact state transitions are given by a starting and a ending state (both stable); *spontaneous* transitions are given by a starting (stable), one or more temporary (instable) and an ending state (stable). Finally, the contact state transitions of the leaf spring in this experiment are shown in Fig. 2. The initial state at the beginning is neutral (N). We are interested in finding the transitions between the following five stable contact states E/E, V/F, E/E, V/F and N. It should be noted that N is a stable state. But in this case, as soon as the spring falls into the slot there is another immediate transition from N to E/E due to the internal tension of the spring. So here N is instable.

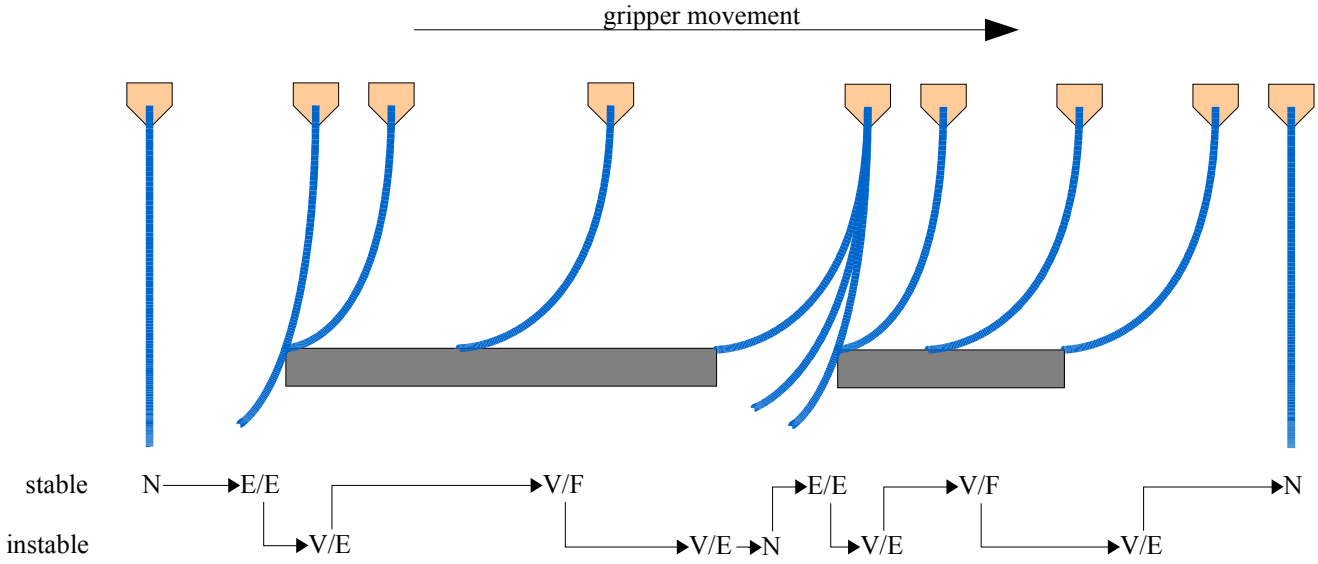


*Fig. 2: Contact state transitions of the leaf spring experiment. The gripper moves from the left to the right. Stable states are denoted in the upper row while instable states are denoted in the lower row.*

**Data Acquisition and Preprocessing**

The impact sound sensor used here in this experiment is able to record signals for three orthogonal axes *x*, *y* and *z*. This means we can use three separate signals to determine the changes in the contact states. Fig. 3 shows the recorded signals for the described experiment on the left. The samples are recorded with a rate of 24kHz per axis. The movement of the robot starts at around 1.2 seconds from the start of the recording. The changes in the contact states occur at about 3.3 (N→E/E), 4.5 (E/E→V/E→V/F), 5.5 (V/F→V/E→N→E/E), 6.1 (E/E→V/E→V/F) seconds. At about 6.7 seconds, the last change (V/F→V/E→N) occurs. It can be seen that all three signals are nearly identical. Unfortunately, the only state change that can be seen directly in all three signals is the one at 5.5 seconds when the spring falls into the slot in the plate and collides with the inner edge. The other changes cannot be detected directly in the signals due to the vibration that is caused by the robot motion.

To eliminate the noise generated by the robot motion, we have recorded a reference signal of exactly the same movement, but without the leaf spring in the gripper. So all vibrations measured by the sensor are generated by the robot. The original and the reference signal of the z-axis are displayed in Fig. 3 on the right. Inspecting the amplitudes of the frequency spectrum over time (Fig. 4, right) we can see that the noise generated by the robot consists of only low frequency signals that are certainly generated by robot motion and are limited to the range of roughly 0 to 400 Hz. On the other hand we can clearly see in the frequency spectrum of the original signal that the impact of the spring with the plate generates a signal that is more or less evenly distributed across the whole frequency range (Fig. 4, left) but with a significantly lower amplitude than the robot signal.

As the noise of the robot motion is limited to the lower end of the spectrum, we can use a highpass filter to eliminate the noise and get a clear signal of the sound generated by the leaf spring. We have used a Chebyshev filter with six poles and a cutoff-frequency of 400 Hz. The filter coefficients can easily be calculated using the method described in [10].

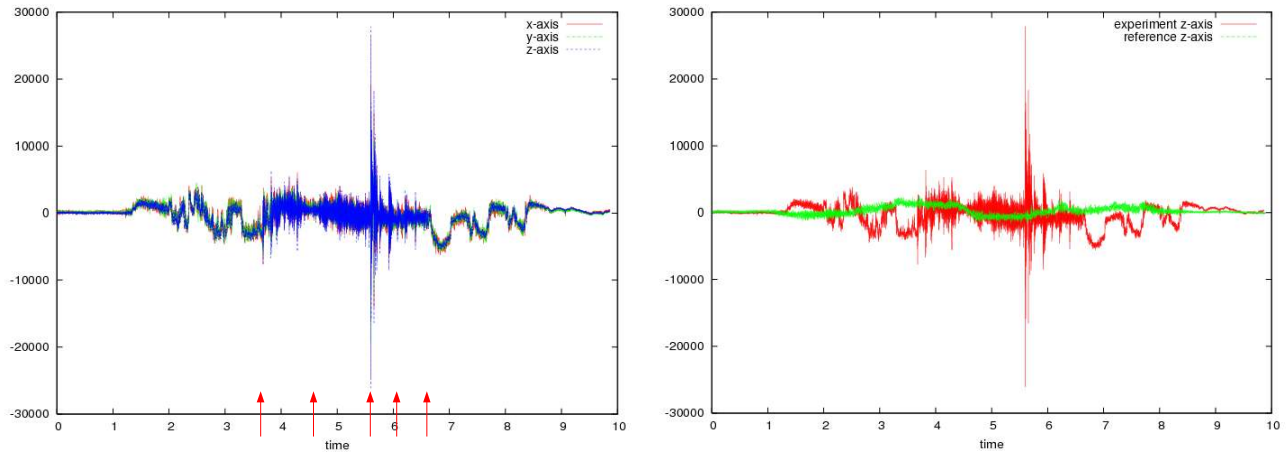


*Fig. 3: Experimental and reference signals of the impact sound sensor. Left: All three experimental signals. Transitions are marked by an arrow. Right: The experimental and the reference signal of the z-axis.*

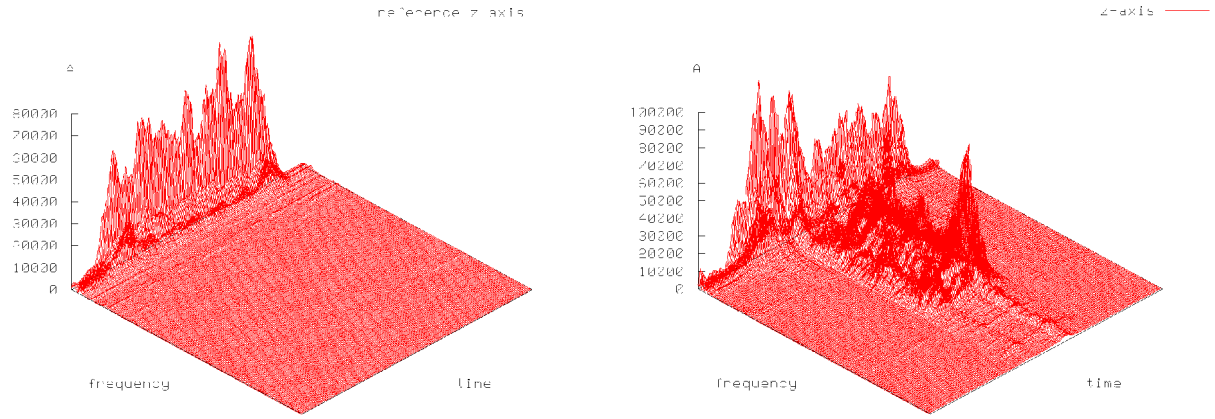


*Fig. 4: Left: Frequency-amplitudes displayed over time of the reference signal. Right: The experimental signal.*

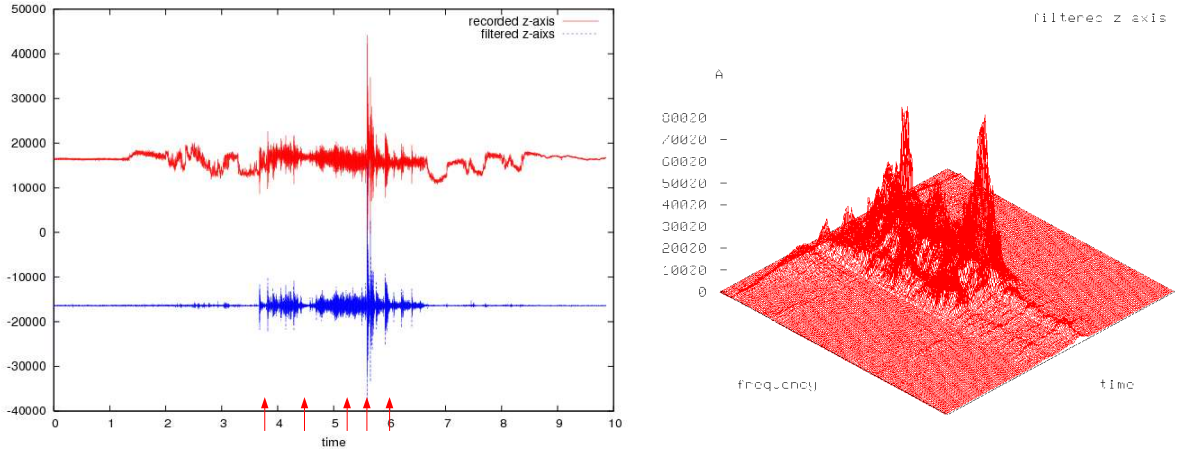


*Fig. 5: Left: Recorded (red, top) and filtered (blue, bottom) signal of the z-axis. Both signals were offset by ±16384 for illustration purposes. Right: Frequency amplitudes of the filtered signal plotted over time.*

The original and the highpass filtered signal of the *z*-axis are shown in Fig. 5 on the left. The frequency spectrum over time is shown on the right. It is noteworthy that the filtered signal bears a close resemblance to a common audio signal.

**Feature extraction**

We will now map the transition sequence to this new signal in order to determine the contact state changes. To achieve this we define a set of features to gain further information from this signal.

A common feature used in the analysis of audio information is the *loudness* $L_X(t)$ of the signal $X$ at time $t$. This is simply the standard deviation of the signal. Let $x_t$ be the $t^{th}$ sample of the signal $X$. Then $L_X(t)$ can be calculated over a window of $n$ samples using

$$L_X(t)=\sqrt{E(X^2)-E(X)^2}=\sqrt{\frac{1}{n}\sum_{i=t-n+1}^{t}x_i^2-\left(\sum_{i=t-n+1}^{t}x_i\right)^2}$$

Besides the loudness of the three signals (one for each axis of the sensor) that correspond to the square root of the variances of the signals, we can calculate the covariance between the three signals as well

$$Cov(X,Y)=E(X\cdot Y)-E(X)\cdot E(Y)$$

In our case we get

$$Cov_{XY}(t)=\frac{1}{n}\sum_{i=t-n+1}^{t}x_i\cdot y_i-\frac{1}{n^2}\sum_{i=t-n+1}^{t}x_i+y_i$$

for a window size of $n$ samples. Note that

$$Cov(X,X)=E(X\cdot X)-E(X)\cdot E(X)=E(X^2)-E(X)^2=L_X^2$$

and that $Cov(X,Y) = Cov(Y,X)$. So we find all these features combined in the covariance matrix $LC(t)$ of the three input signals:

$$LC(t)=\begin{pmatrix} L_X(t)^2 & Cov_{XY}(t) & Cov_{XZ}(t) \\ Cov_{XY}(t) & L_Y(t)^2 & Cov_{YZ}(t) \\ Cov_{XZ}(t) & Cov_{YZ}(t) & L_Z(t)^2 \end{pmatrix}$$

The recorded signal is processed in intervals of $n$ samples length. While processing, subsequent intervals overlap over a length of $m$ percent. Fig. 6 shows the loudness (top) and the covariance (bottom) of the three signals for $n$=1024 and $m$=0.85.

The one state transition that can be clearly seen in the diagramms is when the spring falls into the slot (V/F→V/E→N→E/E) . Due to the internal tension the spring snaps to right edge of the slot as soon as the tip of the spring has passed the left edge (V/E contact). This generates a strong impulse that can be seen as a high peak in the loudnesses and covariances of all three channels.

The first transition can be seen as well. As soon as the spring is moved along the edge the resulting friction creates an amplitude that can be detected using a simple treshold value.

The other transitions are much harder to see. The friction of the spring with the plate results in noise that can be seen in both loudness and covariance. But no definite assertions about the type of change of the contact states can be made here. In analogy it can be deduced that the last transition from V/F to N (via V/E) results in an end of the friction. In this case the loudness and covariance fall back close to zero.

It can be seen that the course of the loudness and the covariance of the three signals is very similar. Thus, there is no difference between examining the loudness or the covariance of the signals here.

As we can only appoint the third transition by looking at the loudness (or the covariance) of a signal with certainty, we introduce another feature. So we pursue the following line of thought:

As long as the spring is in the neutral state N the only vibration induced into the spring (and into the sensor) is the one generated by the robot. But as soon as there is a contact with the environment the resulting friction will generate a second signal, as we have seen in Fig. 5. Because this signal is composed of significantly higher frequencies it will shift the mean frequency of the frequency spectrum of the whole signal.

We can transform our input signals to the frequency spectrum using the fast fourier transformation (*FFT*) and then define the mean frequency of the frequency spectrum $M_f(t)$ over the time as

$$M_f(t)=\frac{\sum_{i=1}^{n/2} f_i\cdot \mathrm{FFT}_{[x_{t-n+1}..x_t]}(f_i)}{\sum_{i=1}^{n/2} \mathrm{FFT}_{[x_{t-n+1}..x_t]}(f_i)}$$

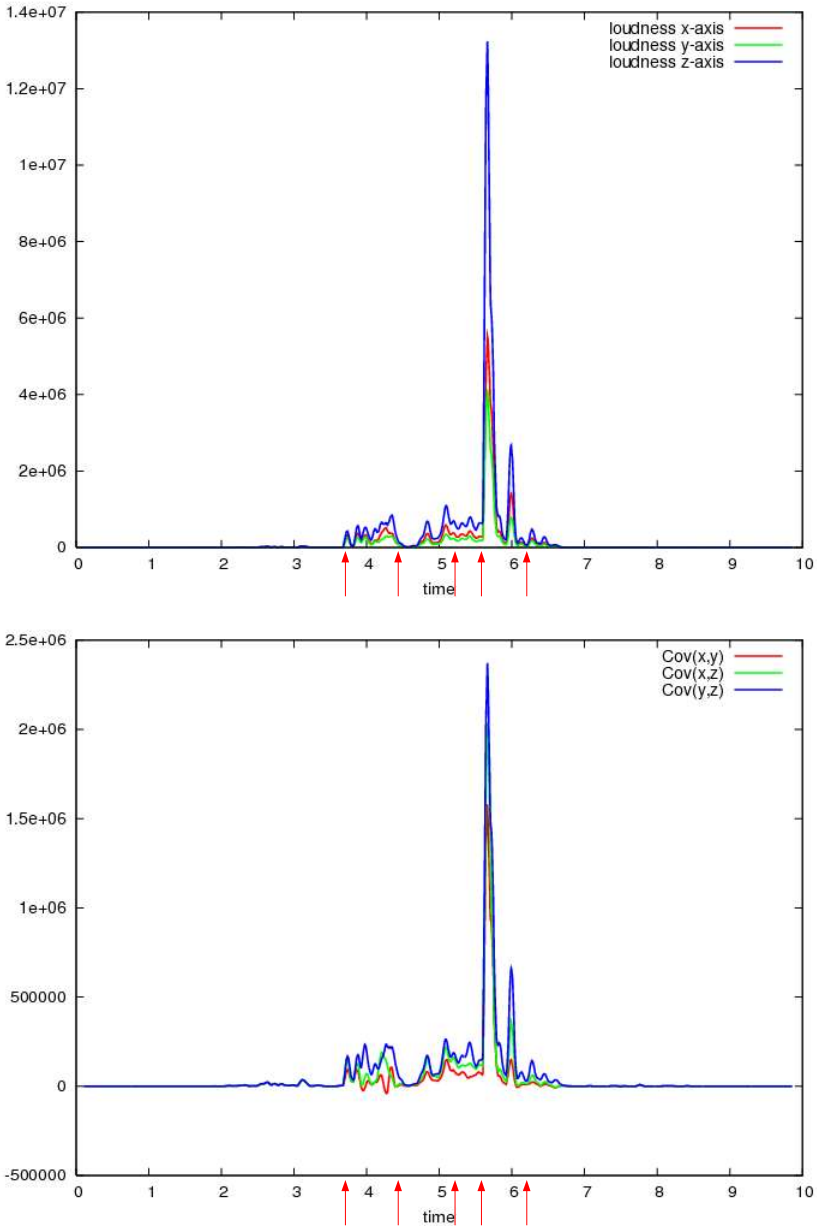


*Fig. 6: Square of loudness (top) and covariances (bottom) of the filtered signal with n=1024 and m=0.85*

It is not practical to make assumptions about the exact nature of the friction noise, that is, its dominating frequencies, amplitudes of certain frequencies and so on, without establishing a detailed physical model of the gripped object and its environment. By looking at the frequency diagram in Fig. 5 on the right we see that the collisions of the spring with the plate results in three dominating frequencies and constant spectrum in the range of 400 to 9000Hz. Although there are visible peaks in the signal, we assume that these are dependent on the speed of the gripper and the materials of the DLO as well as the objects in the environment.

We do not use the highpass filtered signal for this computation as the noise generated by the robot is of interest here. If we used the filtered signal, the mean frequency would be very close to the middle of the spectrum as long as there is no contact with the enviroment. The signal will consist of noise evenly distributed across the whole spectrum with a general low amplitude. When there is a contact, the resulting signal will still mainly consist of noise, even with a higher general amplitude. The mean frequency will still be close to the middle though. If we include the robot noise in this computation, this results in a low value for the mean frequencies as long as there is no contact with the environment (the robot noise will dominate over the white noise). As soon as a contact occurs the increased amplitude of the contact noise will shift the mean frequency more towards the middle of the spectrum.

The frequency spectrum over the samples from a given interval $FFT_{Interval}(f_i)$ is calculated using the well known FFTW library [12] for the fast fourier transform combined with a 4-term Blackman-Harris window [4] that is used to minimize

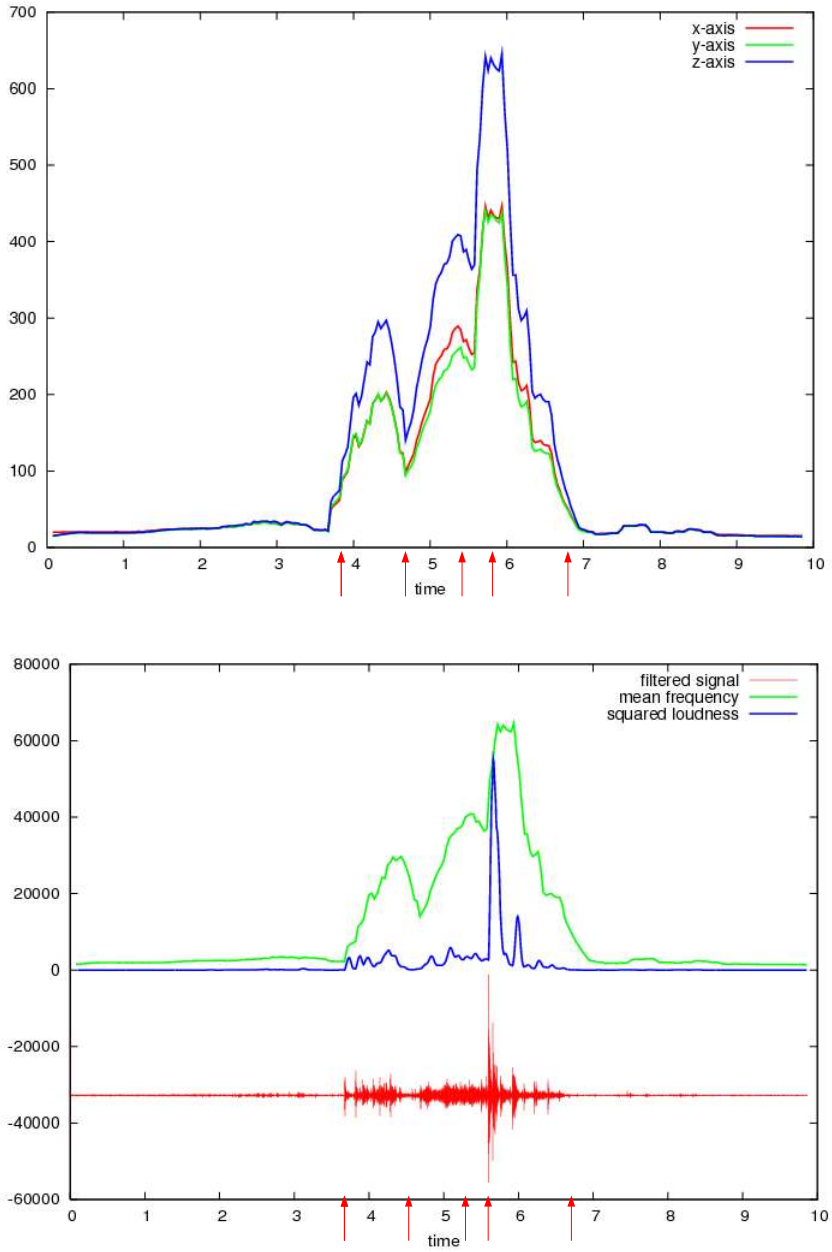


*Fig. 7: Top: Mean frequencies of recorded signals with n=1024, m=15%. Bottom: Filtered signal, corresponding mean frequency and loudness of the z-axis. The filtered signal is offset by -16384, the mean frequency and loudness are scaled by 200 and 1/100 respectively. Transitions are marked by an arrow in both diagrams.*

redundancy in the spectrums. Once more, the recorded signal is processed in intervals of *n* samples length with an overlap of *m* percent of samples from the last calculation.

The mean frequencies of all three signals are displayed in Fig. 7 at the top. Because we use the unfiltered original signal for this computation, we already see a mean of about 20 Hz even if the robot is not moving at all (from 0 to 2 seconds). This vibration may be caused by the robot servo loop. During a simple movement without any external contact this frequency rises to 30 Hz (3 seconds). At the point of the first transition from N to E/E the mean starts to rise for about one second up to 300 Hz (from 3.5 to 4.5 seconds). This is a result of the induced friction from the edge into the spring. The more the gripper moves to the right, the stronger the friction signal and the stronger the shift in frequency. When the tip of the spring reaches the edge of the plate and snaps to the face of the plate (E/E→V/E→V/F) the tension shortly drops which results in reduction of the mean frequency (4.5 seconds). During the movement of the tip along the plate the tension gradually rebuilds. The transition from V/E to N can be seen for a short time, when the mean drops once more. Because N itself is not stable in this case the spring directly enters the next state E/E when it snaps against the right edge of the slot. The impulse generated here has got a high amplitude that reveals itself in a steep rise of the mean frequency (5.6 seconds). While the gripper moves on, the tension of the spring reaches the same level once more. So does the mean. The following

transition E/E to V/F (via V/E) can not be seen here (at about 6.2 seconds). We assume if the robot stopped immediately for a short time after recognizing the contact, this would allow the spring to calm down. In this case the mean would decrease to base frequency of the robot, 20 Hz. As soon as the robot continues the movement, we could recognize the next transition the same way we recognized the first E/E→V/E→V/F transition. Finally, after the last transition (6.9 seconds), there is no more contact between the spring and its environment and the mean falls back down to the value of the robot motion.

Recapitulating, we display the filtered signal as well as its loudness and mean frequency once more in Fig. 7 at the bottom. Using the information from the two described features, we can build an algorithm to detect the contact state transitions. The detection itself can be realized with simple treshold values. When the contact state transitions are identified we can theoretically determine if one of the features (usually the mean frequency) will rise or fall. It is necessary to stop the robot after each transition to reduce the overlap of multiple transitions. As the form and amplitude of the input signals is dependant on multiple factors such as robot speed and material attributes it is hard to define an absolute treshold value for the algorithm without risking false alarms or overlooking a transition.

**Validation of the theory**

We have conducted two more experiments to support our theory: In a second experiment the leaf spring is moved into simulatenous contact with two separate edges (E/E contact). Because the second edge lies about 5 cm under the first edge, the corresponding contacts do not occur at the same time. Instead the second contact is delayed by roughly 1.5 seconds to the first (see Fig. 8, right). The chain of contact state transitions is given as $N \rightarrow E/E_1 \rightarrow E/E_1 \vee E/E_2$ where $E_1$ denotes the first edge of the obstacle, while $E_2$ denotes the second edge.

The results are displayed in Fig. 8 on the right. Once again, we can clearly see the both E/E contacts as peaks in both the loudness as well as the mean frequency at about 3.1 and 4.8 seconds, although the second contact is less pronounced than the first. This is due to the absorbtion of the vibration of the spring by the first contact point. We cannot see any friction here, as the movement of the spring along the edge between the first and second contact is minimal.

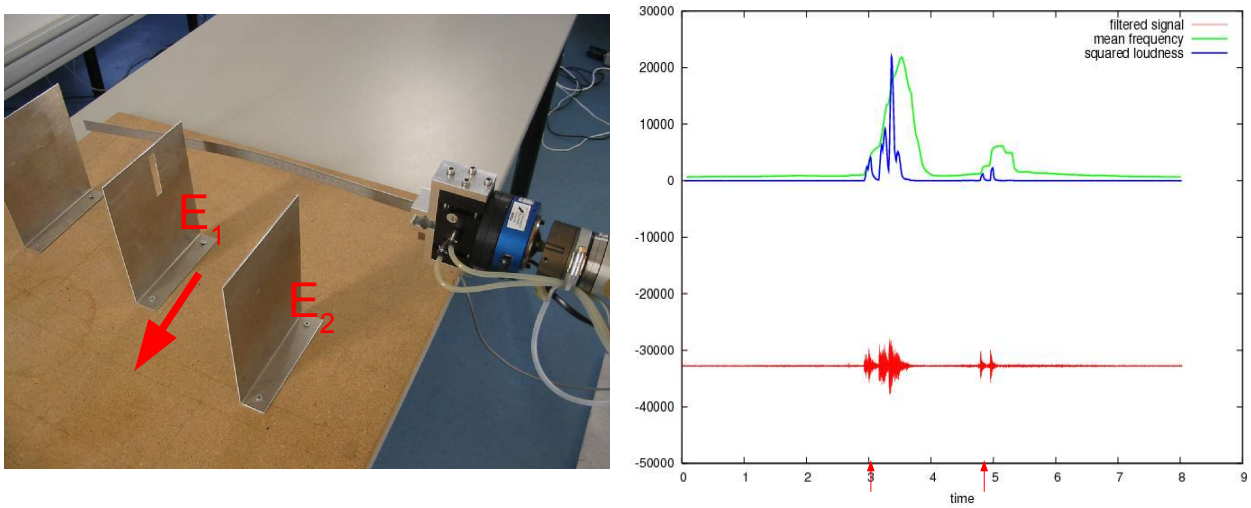


*Fig. 8: Left: Setup of the second experiment. The spring leaf is moved into contact with the first edge $E_1$ then with $E_2$. Right: Filtered signal, corresponding mean frequency and loudness of the z-axis. The filtered signal is offset by -16384, the mean frequency and loudness are scaled by 50 and 1/100 respectively. Transitions are marked by an arrow.*

The third experiment moves the spring straight into a plate, then stops and slowly moves the coiled spring along the plate to the right until if falls of the plate. After this the now extented spring is moved back to the left until it touches the right edge of the plate. The chain of contact state transitions is given as N→V/F→V/E→N→E/E. The setup of this experiment is shown in Fig. 9 on the left, the measured signal and its features are depicted on the right.

Once more, we can clearly see the first transition from N to V/F at 4.1 seconds. Here more than one peak is generated during the contact. This is because the spring snapped to the left due to the high tension when the spring touched the plate nearly normal to the plate. We can determine the period of time from 6 to 7.43 seconds when the robot stopped before commencing the next movement to the right. This movement generates a vibration in the spring that we can clearly see in the diagram. When the spring falls off the plate at 10.8 seconds (V/F→N) the vibration stops. Once again the robot stands still for another 3 seconds to allow the spring to reduce its vibration. Then it is moved back to the right until the next transition (N→E/E) is detected. This contact also generates more than one peak. Once more this is a result of the very low vibration of the spring that was induced when the sping fell of the plate and has not died away completely yet.

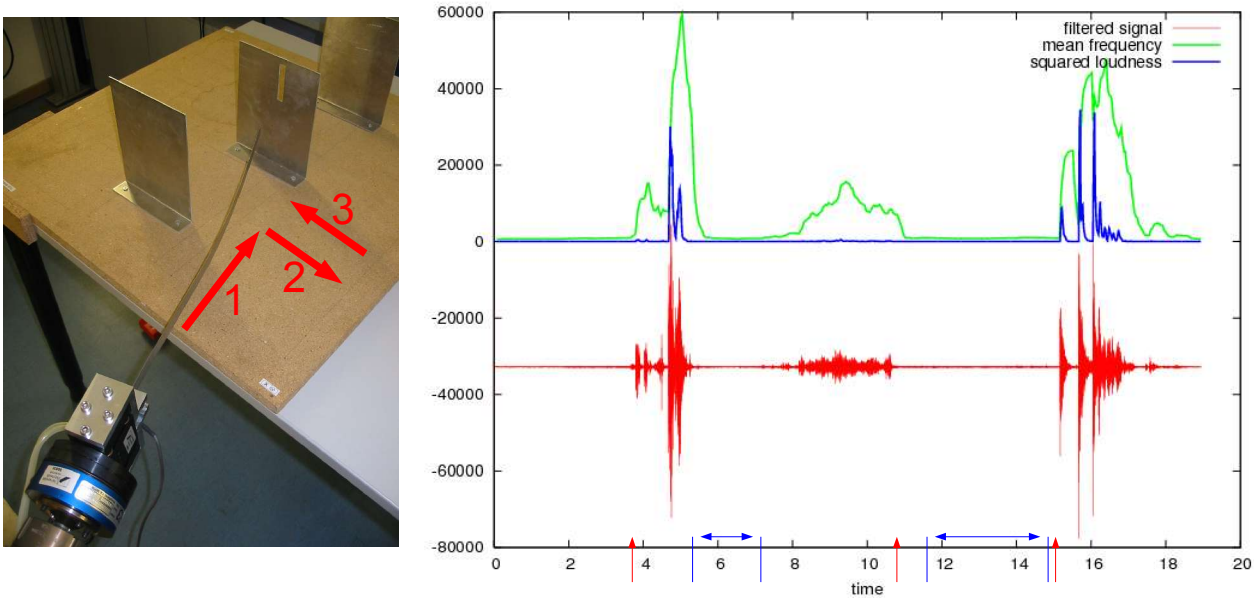


*Fig. 9: Left: Setup of the third experiment. The leaf spring is moved in the three directions until either a contact is established (direction 1 and 3) or is lost (direction 2). Right: Filtered signal, corresponding mean frequency and loudness of the z-axis. The filtered signal is offset by -16384, the mean frequency and loudness are scaled by 50 and 1/1000 respectively. Transitions are marked by a red arrow while intervals without robot motion are marked in blue.*

**Conclusion**

If we compare the two features we see that by using the loudness or the covariance of an acoustic signal we can only determine contacts that generate strong peaks. These usually occur through hard contacts when the object held in the gripper of the robot touches objects in the environment with significant force. Using the mean frequency as a second feature we can determine the state of the spring quite exactly and see other state transitions as well.

Unfortunately the desired gain of information by measuring more than one axis failed to appear. The recorded signals of the three axes are nearly identical. It remains to be seen whether this holds for all directions of movement and all possible contact state transitions.


[1] Abegg F., Remde A., and Henrich D.: "Force- and vision-based detection of contact state transitions". In: Robot manipulation of deformable objects, Henrich D., Wörn H. (Eds.), Springer-Verlag, London, 2000, ISBN: 1-85233-250-6, pp. 111-134

[2] Desai, R. S. and Volz, R. A.: „Identification and Verification of Termination Conditions in Fine Motion in Presence of Sensor Errors and Geometric Uncertainties". In: 1989 IEEE Int. Conf. on Robotics and Automation (ICRA'89)

[3] Dutre, S., Bruyninckx, H. and De Schutter, J.: „Contact Identification and Monitoring Based on Energy". In: 1996 IEEE Int. Conf. on Robotics and Automation (ICRA'96)

[4] Harris, F. J.: "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform". In: Proceedings of the IEEE, Vol. 66, No. 1, January 1978

[5] Henrich D.,Ogasawa T., Wörn H.: "Manipulating deformable linear objects: Contact states and point contacts". In Porc. 1999 IEEE International Symposium on Assembly and Task Planning (ISATP'99), Porto, Portugal, July 21-24, 1999

[6] Mosemann, H., Raue A., and Wahl, F.: „Classification and Recognition of Contact States for Force Guided Assembly". In: 1998 IEEE Int. Conf. on Systems, Man and Cybernetics

[7] Remde A., Henrich D., Wörn H.: "Manipulating deformable linear objects: Contact state transitions and transition conditions". 1999 IEEE/RSI International Conference on Intelligent Robots and Systems (IROS99), Kyongju, Korea, October 17-21, 1999

[8] Schlechter A., Henrich D.: "Manipulating Deformable Linear Objects: Characteristics in Force Signals for Detecting Contact State Transitions". 10th International Conference on Advanced Robotics (ICAR 2001), Budapest, 22.-25. August 2001

[9] Schlechter A., Henrich D.: "Discontinuity detection for force-based manipulation", In 2006 IEEE Int. Conf. on Robotics and Automation (ICRA'06), Orlando, Florida, USA, May 15.-19.2006

[10] Smith, Steven: "The Scientist and Engineer's Guide to Digital Signal Processing", www.dspguide.com

[11] Wackerly D., Mendenhall W., Scheaffer R.: "Mathematical statistics with applications", Duxburry Advanced Series

[12] www.fftw.org

[13] Xiao, Jing: „Automatic Determination of Topological Contacts in the Presence of Sensing Uncertainties". In 1993 IEEE Int. Conf. on Robotics and Automation (ICRA'93)