POLITECNICO DI TORINO

Master degree course in Electronic Engineering

Master Thesis

Design of a force sensor for an haptic and touch interface

Supervisor
Maurizio Zamboni

Candidate
Mattia Primavera

February 2019

To my grandparents Egeo, Gina, Piercarlo and Rosanna

Acknowledgements

I would like to thanks all the people that supported me from the beginning of my university career : my parents that always spur me to do better and help me in realising my dream in every possible manner. My girlfriend Giulia for helping me and giving me advices in those situation where I was in trouble and for always supporting my decisions. My sister Allegra for motivating me and always be by my side when it was necessary. My friends and team mates Calogero, Mattia and Simone for sharing with me lot of nights studying and thinking. My childhood friends for always being a support. Professor Maurizio Zamboni for helping me in these 5 years of university providing me a solution for every problem I encounter. Trama Engineering, all the colleagues for helping me in my research. Flavio Cerutti for always helping me giving advices and stimulating my imagination.

Contents

List of Figures

Chapter 1

Introduction

The aim of this thesis is to design an innovative system that can monitor the force applied by the user on a touch surface in order that the system is able to understand its intentions. This need derives from the development of touch interface that historically provide informations in a two dimensional world, however the user is used to touch, press an object and for this reason a third dimension is required.
This technological trend has been practised by all the major tech company that start introducing force detection sensors on their touch ecosystem allowing the introduction of haptic feedback systems. Haptic is the science of touch and of how people can interact with the world. Nowadays it is part of the process that tries to reduce the gap between human and machine interface. In fact to improve the capability of an object to transmit informations sight and hearing are no more sufficient, there are feelings that can be felt only with touch. For example the replacement of a mechanical button with a touch one is possible only with an haptic feedback that is the better solution to warn the user the operation has been correctly done. This allow the introduction of this technology also in automotive cockpit where touch technology has been rejected for years for security. However with the integration of touch technology with the haptic one even field where the safety is important can be reached, such as car steering wheels. For this reason luxury car makers have started working on it in order to begin a process of modernization.
Understanding the incoming role that this technology will have in the future, I decided to work on a technology that could help the integration of the touch and haptic technology allowing new possibilities and new field of applications. So I focus my attention on the touch force measurement that is a needful part of the system and that is the reason why touch and haptic can be introduced in complex fields. The reason why touch has not been introduced in automotive field was the fact that the probability of an unwanted touch was too high. However the introduction of the force detection combined with the touch technology allow to produce a reliable system.
Trama Engineering, the company where I carried out my thesis, needed a more precise force sensor to be integrated with their haptic actuator. Studying the haptic actuator and its behaviour came the idea to exploit the same physical mechanism to measure the force applied. This particular solution allow to measure the force without the introduction of a proper sensor but only using the actuator. For this reason it can be considered a sensorless system, more efficient and faster than the others solutions. After the study of the physical

feasibility, the development of hardware and firmware enable the realization of a correctly working system. Finally a testing campaign is performed to verify and certify the behaviour in various conditions. This allowed me to understand all the phases of the development of a product: from the conception of an idea, to the design phase, the realization and the verification of its working.

Chapter 2

Overview on touch technology

2.1 Touch Sensing Technology

In recent years Touch Technology has become more and more important and commonplace on every types of product on which traditionally a mechanical button has been used. This is due to his flexibility and user friendliness. The user in this way can interact directly with the display rather than with a touchpad or a mouse and so there is nothing more natural than touch a thing directly on screen, as children do. Touch interfaces allow to reduce the mechanical parts enabling a cut in the manufacturing cost and time, a reduction of failure risk and other important benefits such as the fact that they are waterproof while the mechanical one can be permanently damaged when water slips on it. However touch solutions can often detect false positive touch events due to the moisture on the sensing surface or can be affected by noise if there is not a sophisticated firmware algorithm[1].
There exist different types of technologies used for the touch sensing each one characterized by its own merits:

- **Resistive Touch** : The resistive touch technology is the most widely used touch technology today due to its low cost of production. It is composed by two resistive and flexible metallic sheets separated by microdots and air.
  There are two types of metallic layers, one called "matrix" with striped electrodes

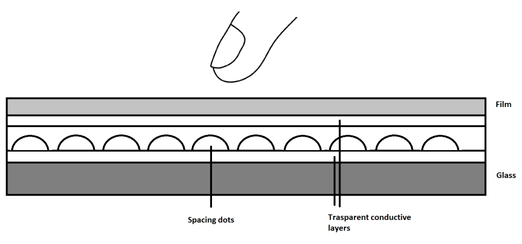


Figure 2.1. Resistive Touch panel

on two surfaces face each others of glass or plastic, the other one called "analogue"

with transparents electrodes without any pattering. When any object contact with the surface of touch screen the two sheets are pressed together and since on these sheets there are horizontal and vertical lines when they are pressed, they get in touch each other allowing a precise localization of the touch. For this reason they have high resolutions providing an accurate touch control.
Since this technology senses the contact with any object it is defined as a "passive" technology. It has also others advantages such as the low power consumption, the low cost of production, tactile feel and the resistance to water and others liquid. However the external film is vulnerable to damage from sharp objects and from scratching.

- **Surface Capacitive** : For this Technology the screen is composed of a transparent conductive coating recovered by a glass protective coating on it. On the four corner of the screen four electrodes are placed and the same voltage is applied to each one, creating an uniform electric field on the panel.

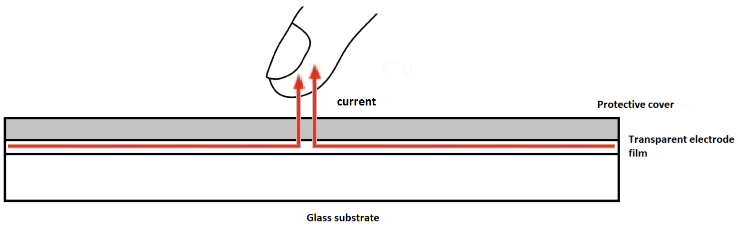


Figure 2.2. Surface Capacitance Touch panel

When the finger touches the screen an electric current will flow from the four electrodes to the finger and measuring the current from each electrode it' s possible to know if there is a finger present and where it has been placed.
The ratio between the four currents provides information of the position. Of course the current of the electrode is inversely proportional to the distance between the electrod placed in the corner and the finger.
This technology provides high resolution, high response speed and can respond to light touch, in fact no force is needed for the detection (but this makes it affected by noise).

- **Projected Capacitive** : This is the most used technology, used on smartphones. A projected capacitive sensor has X and Y electrodes that are placed on differents glasses but intersecting each other in a matrix. When the finger touches the screen since it is conductive a capacitance coupling occurs between it and the electrodes. In this way the capacitance between X and Y electrodes changes and the system can detect it. So this technique exploit the conductivity of humans skin[2].
Between the advantages the excellent image clarity, the resistance to liquid contaminants and the possibility of Multi-touch. The bigger disadvantage is the affection by noise.

- **Surface Acoustic Wave** : This particular technology utilize a series of piezoelectric transducer and receivers on the sides of the touch area. The transducer creates

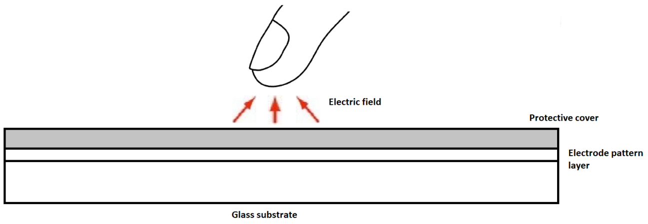


Figure 2.3. Projected Capacitive Touch panel

ultrasonic waves that travels on the surface. When the user finger is placed on the surface a portion of the waves are absorbed and so the receiver can notice it and localize the touch point. This solution provides and excellent image clarity and an

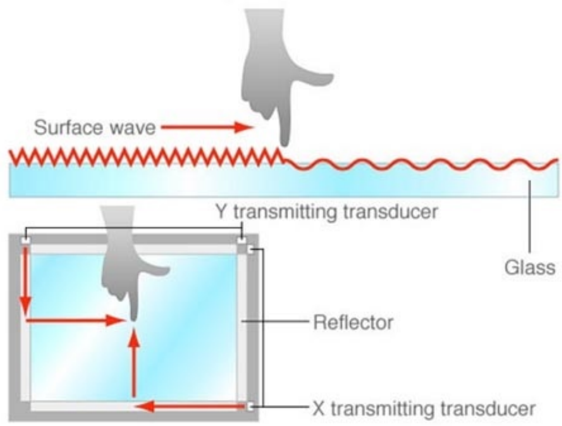


Figure 2.4. SAW Touch panel, credit to "*Acoustic Based Sensing Tech and Its Application for Novel Touchscreen Concept*", R.-Y.Fong, H.Qin

high "touch-life" but it has problems if there are some contaminants on the screen that would create non-touch zones or if the object are too hard (since it would not absorbe waves). Also water and other liquids falling on the surface can create false trigger events.

- **Infrared Touch** : this solution is made possible thanks to Infrared emitters and receivers placed on the corners of the screen that create an invisible grid of light beam. This provides the best image quality solution since the measuring system is above the screen and for the same reason its touch life is unlimited. The detection happen because the finger blocks the light and so the correspondant X and Y receivers would not receive the signal.

However for field like automotive a touch sensing technique it' s not sufficient cause the user can activate accidentally a function or a touch button only by brushing on it. But if the user is travelling at 130 km/h and the function activated is the electronic handbrake this can be very dangerous. For this reason the system requires also the force applied to the surface.

2.2 Force Sensing Technology

Conventional touch displays can detect two dimensional touch information, so the position of the contact along the X-axis and the Y-axis but don't provide information about the Z-axis. Some touch sensing technology such as the capacitive one provide an information about how close an object is from the touch surface, but this happens within a specific threshold. A system with only the touch sensing technology limits the user interfacing possibilities and the build of a more complex but real interface. For this reason today in almost every touch screen a Pressure Force is present to reduce accidental touches and provide a more precise experience.
Also for the Force Sensing Technique there are different solutions:

- **Resistive Force Sensor** : It' s a sensor made of a particular material whose resistance changes when a force or mechanical stress is applied. It is composed by two sheets of flexible, conductive polymer ink together that can easily be printed.

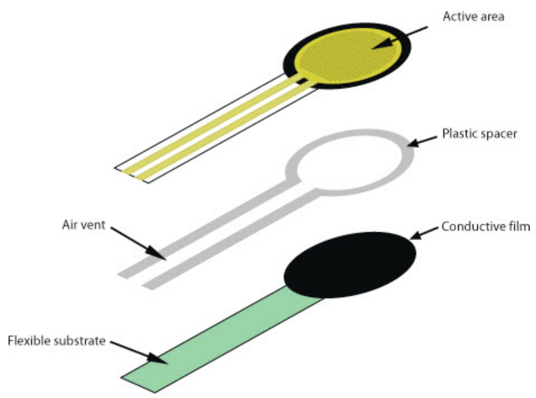


Figure 2.5. Resistive Force Sensor structure, credit to "*www.elprocus.com/force-sensing-resistor-technology*"

When a force is applied the sensing film represented in the picture by the yellow active area get in touch with the below conductive film changing in this way the resistance of the film. This sensor requires a very simple interface that should only read these variations and transform them in a Force information.
The advantages are not only the low cost of production and a good shock resistance but the small size. In fact despite the other sensors this has a thickness that is typically less than 500 μm, and this is a key point in a world where the scaling of dimensions is becoming more essential. However it is characterized by a low precision.

- **Piezoelectric Load Cells** : this sensor employs the piezoelectric effect for the measurement of the force. When a force is applied the sensor product an high impedance electrical signal that is then modified by an amplifier. It has a very fast response and can measure also quasi-static forces.

- **Capacitive Force Sensor** : This sensor is made of a material whose capacitance changes when a force is applied. Their behaviour is very simple: the two plates of the touch screen works as two plates of a capacitor. Since the capacitance C is inversely proportional to the distance between the two plates, if the distance became smaller the C grown and so it can be detected. In the structure are also present some

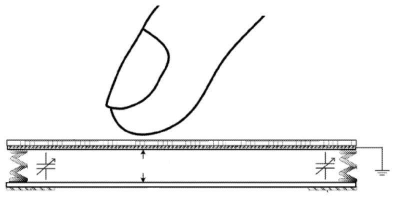


Figure 2.6. Capacitive Force sensor

elastic components that keep constant the distance between the plates when no force is applied. It is important that the material is linearly elastic in order to simplify the computation of the Force. Since for the Hooks Law $F = k\ x$ it' s possible to write the following equation for the capacitance: where $d_{nominal}$ is the distance where no force is applied, while $d_{nominal} - \frac{F}{k}$ is the distance between the plates when a force is present. Rearranged it we obtain:

$$C = \frac{\epsilon\ Area\ d_{nominal} + \epsilon\ Area\ \frac{F}{k}}{d_{nominal}^2 - \frac{F_{finger}^2}{k^2}}$$

If the deformation x is very small it is possible to approximate the equation :

$$C \simeq \frac{\epsilon\ Area\ d_{nominal} + \epsilon\ Area\ \frac{F}{k}}{d_{nominal}^2} = C_{nominal} + \frac{\epsilon\ Area\ F}{k\ d_{nominal}^2}$$

Since the term $\frac{\epsilon\ Area}{k}$ is a constant it is called ϕ.

$$F = \frac{\Delta C}{\phi}$$

So this is the final expression for the Force applied through the measure of the capacitance. This solution is very simple but has an important disadvantage, the noise affection.

2.3 Haptic Feedback Technology

All the technologies seen before provide a good user experience but in particular cases their expressiveness is limited, because the information that a device would share with us are normally shared through visual or auditive interfaces, such as displays or speakers. But nowadays this is no more sufficient. Haptics open new scenarios for the improvement of the relations between human and machine enhancing more realistic virtual interactions.
It allows to refine on these areas where the sense of touch is critical, such as a remote surgery or teleoperating a robot in the space avoiding a dangerous human space walk, or where touch can replace visual and auditory task, like in mobile communication and gaming [3].In fact it is very difficult to imagine a world without touch because it is an integral part of our lives. Without it humans would have a great difficulty manipulating objects, discovering the material properties. Computers, smartphones and tablets have already input sensors that provide the informations to the machine but their outputs are not sufficient and for this reason the development of haptic technologies would help this aspect.

2.3.1 Human Haptic Perception

Humans have at least five senses that help perceiving environmental information. Touch is different from the other senses because it is not located in a particular part of the body, in fact it is present over the whole skin. It is a very complex system that helps human to respond to external physical environment and it is also very fast, since the touch is 20 time faster than vision according to Heller and Schiff[4]. The sense of touch is divided into two different sense modalities: kinesthetic and tactile.

- Kinesthetic sense is perceived by muscles and tendons and provide sensations about forces and torques.
- Tactile sense instead provide informations about pressure, vibrations and temperature thanks to particular sensory and organs called "mechanoreceptors" situated in the skin.

There are differents types of mechanoreceptors situated in differents parts of the body and specialized to respond to a specific haptic stimulus. The Meissner corpuscles responds to light touch and low frequency vibrations (up to 50 Hz). Pacinian corpuscles provides informations about transient touches and respond to a wide range of frequencies (40-400 Hz with a maximum sensitivity at 250 Hz). Merkel disks detect edges and spatial features in a range up to 15 Hz.
Experiments performed by the professor Bolanowski show that the vibrotactile information is provided by Pacinian corpuscles at medium frequencies and by Non-Pacinian corpuscles at low frequencies, indicating that the two processes are separated [5].

Figure 2.8 shows that the touch is sensible to vibrations up to 1 kHz, that is the band of the touch sense, with a peak at 250 Hz. This means that a vibration at 250 Hz can have a lower intensity to be sensed with respect to a vibration at 10 Hz. With a better comprehension of human touch it is so possible to understand how the haptic devices work

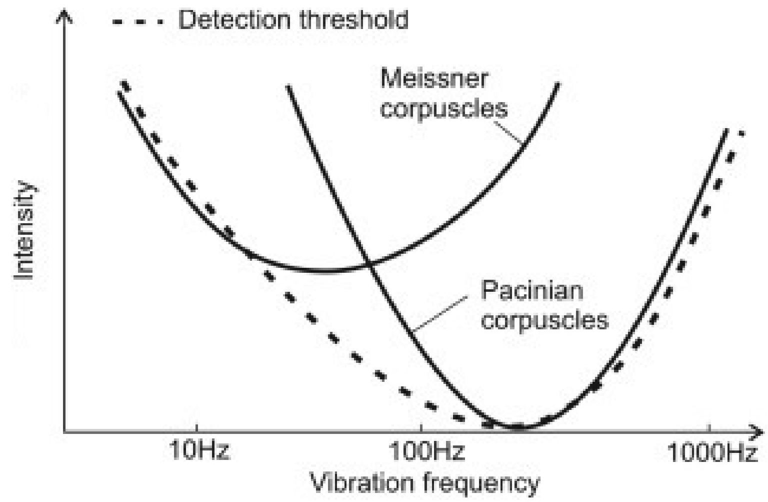


Figure 2.7. Threshold of detectability of vibrotactile stimulus, credit to "*Conn's Translational Neuroscience*", P.M. Conn

and how they act on our body.

2.3.2 Haptic Feedback Actuators

There are differents types of Haptic Feedback Technology to provide the sense of touch to user:

- **Linear Resonant Actuator LRA** : the Linear Resonat Actuator is composed by a magnet surrounded by a coil and attached to a spring. When a AC voltage is applied to the coil,it generates an electromagnetic field that cause the moving back and forth, resembling the behaviour of a speaker producing sound. When the coil is driven to the resonant frequency of the spring the entire actuator vibrates. In order to feel a perceptible force the coil must work at the resonant frequency and this means that the LRA can work only at a specified frequency and can not range,unless the coil is not changed. This implies also that the LRA needs a certain time to start working, that is approximately 10 ms. The stop time instead can be a lot larger, about 300 ms , due to the kinetic energy stored in the spring. However a new breaking system has been developed to stop it vibrating and it let the mass stopping within 10 ms.
  The Apple Taptic Engine, responsible of the haptic feedback in each Apple device since 2016 with the release of iPhone 7, is based on a LRA.

- **Eccentric Rotating Mass ERM** : This actuator has been the most used since 1990 offering an haptic feedback through vibrations. It is composed by a DC motor at whose extremity is fixed an off-centered mass. When the motor rotates the force of the mass became asymmetric causing a centrifugal force that provides a vibration feeling. They are used in mobile and cell phone industry to alert the user about calls or text messages but also in game console like PlayStation with its DualShock controllers. It is a very simple actuator which has a very cost of production, however

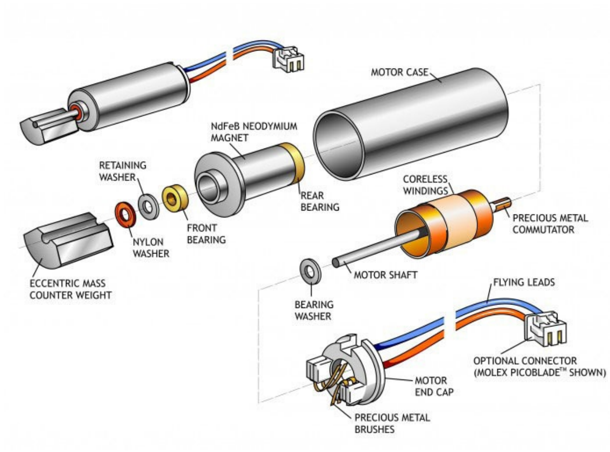


Figure 2.8. Structure of a ERM, credit to "*Tactile actuator technologies*", V.Huotari

the market of actuator is moving away from it because its dimensions that are not so small, because of its lack of precision and because it requires a lot of power. It requires also a lot of time to start producing a vibration, 30-50 ms that compared to others actuators is a very large time.

- **Piezoelectric Actuators** : Piezoelectric actuators are made of a special type of ceramic that contracts when an electrical charge is applied. If a differential voltage is applied to the piezo it starts bending creating a vibration. Despite the ERM in this case it is possible to control the vibration through the voltage and frequency of the signal applied. In this case since the frequency of the vibration is not bound to

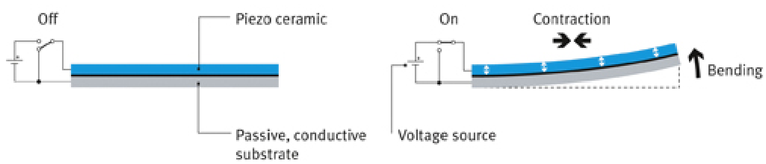


Figure 2.9. Diagram of a Piezoelectric actuator ,credits to "*Festo AG&Co*"

the resonant frequency of the mechanical system it is possible to chose in large range the frequency wanted[6]. So for what concern the precision of the vibration Piezos are more precise than both LRA and ERM. However they are very fragile due to the mechanical properties of the ceramic compound and requires an higher power respect to the previous actuators.

- **Niceclick** : The Niceclick is a new Haptic Feedback actuator developed in 2017 by an italian company, Trama Engineering. Its structure is so composed.It has a metal body of 11SMnPb30 and inside a copper coil 700 μm high. An O-Ring, 120 μm and 20 ShA of hardness, is placed on the superior edge of the coil and it has the function to increase the force feedback acting like a spring. Over it there is a metallic pad which is the moving part that will be fixed to the touch surface. A screw keep the pad fixed to the metallic body and a rubber disk is placed between them to minimize the horizontal movements.

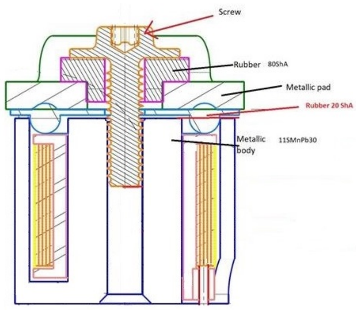


Figure 2.10. Niceclick section, credit to "*Trama Engineering*"

It is different from each of the previous actuators and it has been conceived because of the need to create a feedback that is different from the usual vibration. In this way the user can feels different feedback based on the information that the device want to tell him. There can be thousands of different feedbacks based no more only on vibration but on pulses. It provides the feedback by exercising a variable magnetic force on the metallic pad that will be attracted downward. Despite others actuators, such the piezos which requires high voltages driving circuits, this one can be managed by simply regulating the current flowing in it. This operation can be done by a microcontroller that through a pwm signal can drive a low side mosfet. This allow the system to be really fast and flexible. Higher the current higher the force practised, higher the acceleration provided to the armature and so the feedback. In this way by managing pwm signal it is possible to create every type of pulse with the desired frequency. An example is reported in the graphs below where a Niceclick feedback has been recorded thanks to the use of an accelerometer.

As is it possible to see in the first graph the pulse starts with an high positive acceleration (downward direction) due to an high driving current. Thanks to the elastic component (the O-Ring) the force accumulated is not all transferred in half a millisecond to the pad that would clash to the body causing mechanical breaks. Instead it allow to extend the pulse because when it is stretched,according to the Hooke's law , it generates a restoring force that give back part of the force energy. Exploiting this phenomenon and the related harmonic motion it is possible to creates the desired force spectrum, in Fig. 2.14. Due to human properties, explained before, all the energy of the force signal is below 1 kHz which is the human touch bandwidth. Different

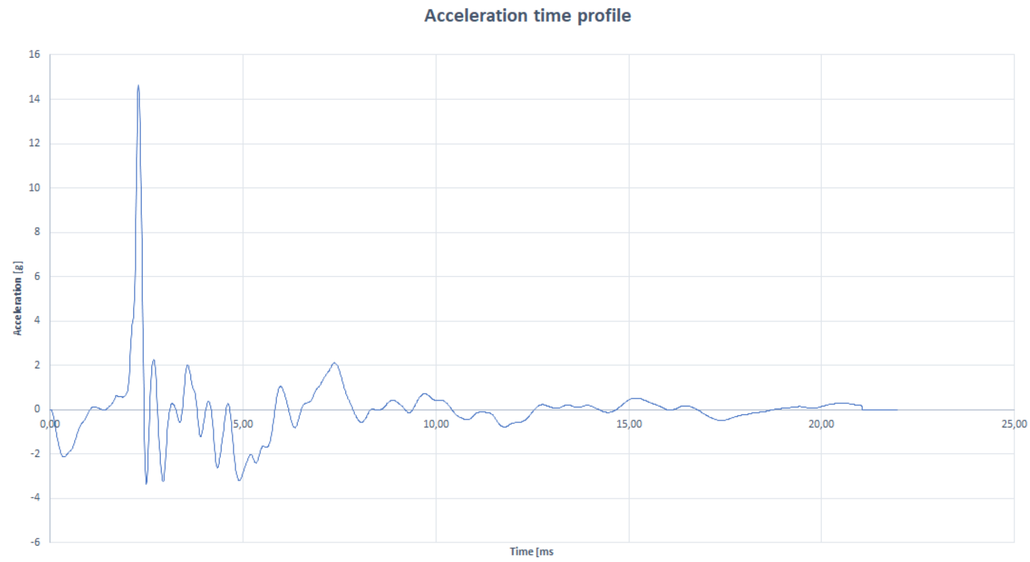


Figure 2.11. Acceleration profile in time domain

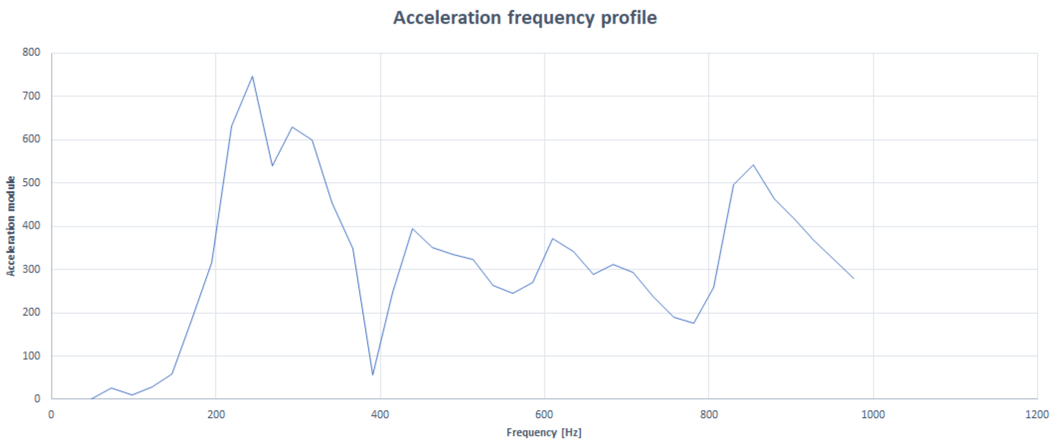


Figure 2.12. Acceleration profile in frequency domain

peaks are placed to various frequency in order to stimulate all the skin receptors and provide a pleasing feeling.

Chapter 3

Force Sensor Design

Despite the others sensor presented before, the work of this thesis concern the design of a new method to measure the force applied to a surface where a Niceclick actuator is planned. This technique allow to exploiting the actuator itself to measure the force. This can be done because the actuator, which is fixed to the movable touch surface, is subjected to the displacement caused by the pressure of the user. Taking advantage of this detail and knowing the model of the actuator the following system has been created.

3.1 Electromagnetism theory and principles

The designed sensor exploit some concept of Electromagnetism, the same that enable the actuator to create a pulse but used as an input information. It is based on the magnetic circuit of a solenoid enclosed into a metal body that allow to hold and direct the magnetic flux into a precise direction.

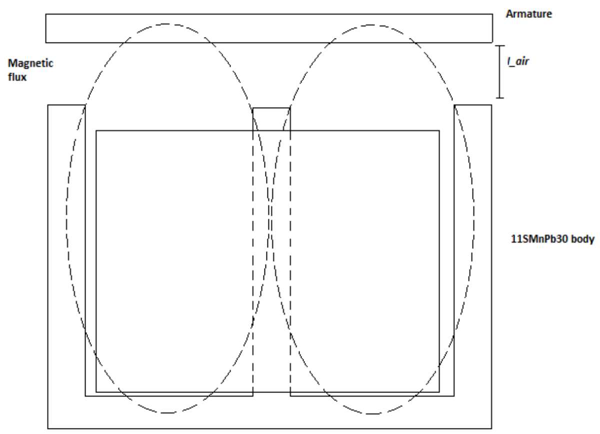


Figure 3.1. Section of the magnetic circuit of the Niceclick

The magnetic flux B of the solenoid can be created when an electric current pass through it, belonging to the Ampère' s Law that in this case can be written as

$$B = \mu H = \mu \frac{N I}{l}$$

Where N the number of the turns, l the lenght of the coil and I the current flowing in it. The SI unit of the magnetic flux is A/m. An important note is the fact that when analysing magnetic loops an exact theoretical computation is not possible and not required, for this reason some approximations are done to simplify the reasoning. For example in the previous formulation the magnetic field is assumed to be homogeneous both in the iron and in the air gap, and so it derives:

$$\oint H \, dl = l \, H = N \, i$$

The soft iron core increase the magnitude of the magnetic flux density and it is taken in account by the term μ that represent the magnetic permeability of a material, in this case iron that is a ferromagnetic material. The magnetic permeability is the ability of a material to support the formation of a magnetic field in itself and so represent the magnetization that a material receives when subjected to a magnetic field. Another parameter often used is the relative permeability μ_R that is the ratio between the magnetic permeability of a specific material with the air one.

$$\mu_r = \frac{\mu}{\mu_0}$$

This a very import parameter because it has a strong dependency with the frequency. This sensor base his properties on another parameter that is the Inductance L of the solenoid. The inductance is a property of conductors according to which any change in the magnetic flux induces an electromagnetic force in the conductor. The inductance of a solenoid can be written as the ration between the flux linkage and the current flowing in the solenoid itself, as reported below:

$$L = \frac{\lambda}{I}$$

From the definition of the MagnetoMotive Force it is possible to find the expression of that describe the magnetic flux Φ and the flux:

$$N \, I = \Phi \, \mathcal{R} \quad \rightarrow \quad \Phi = \frac{N \, I}{\mathcal{R}}$$

$$\lambda = N \, \Phi = \frac{N^2 \, I}{\mathcal{R}}$$

The inductance of the coil is so :

$$L = \frac{\lambda}{I} = \frac{N^2}{\mathcal{R}}$$

The term $\mathcal{R}$ represents the Reluctance of the circuit and express the difficulty to induce a magnetic field in it. It can be expressed as the ratio between the Magnetomotive Force and the flux:

$$\mathcal{R} = \frac{mmf}{\Phi}$$

$$\Phi = \int B \ dA$$

$$mmf = \int H \ dl$$

Assuming the uniformity of the flux density of our circuit in the air gap between the core and the plate, a soft iron core that limitates the saturation effect and a no leakage on flux[7] it is possible to simplify it with:

$$\mathcal{R} = \frac{mmf}{\Phi} = \frac{H_m \ l_m}{B_m \ A_m} = \frac{l}{\mu \ A}$$

where l is the lenght of the circuit, A the cross-sectional area of the circuit and μ the permeability of the material. Since the magnetic sensor circuit has two different materials, lengths and areas, the resultant Reluctance is defined by:

$$\mathcal{R} = \mathcal{R}_{airgap} \ + \mathcal{R}_{iron}$$

$$\mathcal{R} = \frac{2 \ l_{air}}{S_{air} \ \mu_0} \ + \ \frac{l_{iron}}{S_{iron} \ \mu_0 \ \mu_r}$$

Since the cross section of the two material are the same it is possible to simplify the expression:

$$S = S_{air} \ = S_{iron}$$

$$\mathcal{R} = \frac{l_{iron} \ + \ 2 \ l_{air} \ \mu_r}{S \ \mu_0 \mu_r}$$

This last expression can be exploited to rewrite the inductance one, that becomes:

$$L = \frac{N^2}{\mathcal{R}} = \frac{N^2 S \ \mu_0 \mu_r}{l_{iron} \ + \ 2 \ l_{air} \ \mu_r}$$

Since the magnetic permeability of the iron used for the fabrication of the surronder body of the coil is a lot higher this of the air $\mu > 1000 \ \mu_0$ it is possible to approximate the total reluctance with the air one (this is reported only for completeness since the equation used is the previous one).

$$\mu_{iron} = \mu_r \mu_0 = 6.3 \ 10^{-3} \ \frac{H}{m} \qquad \mu_0 = 1.256 \ 10^{-6} \ \frac{H}{m}$$

$$\mathcal{R} = \mathcal{R}_{air} \quad \rightarrow \quad L = \frac{N^2 \ S \ \mu_0}{2 \ l_{air}}$$

This equation is very useful because allow us to express the inductance of the solenoid as function of the distance between the body and the plate of the actuator $L(l_{air})$.
This means that if there is a variation of the air gap due to an external force applied to the armature, the inductance will vary:

$$\Delta F_{finger} \rightarrow \ \Delta l_{air} \rightarrow \ \Delta \mathcal{R} \rightarrow \ \Delta L$$

3.2 Hardware Design

Studying the behaviour of the reluctance and thus this of the solenoid inductance a circuit that can exploit it has been developed. It started from the general idea that varying the value of the inductance also the energy stored in the solenoid changes. The circuit so must be able to convert this energy variation into a varying signal that can be easily analysed by a microcontroller.

3.2.1 Charging phase

The first step necessary to observe inductances variation is the excitation of the solenoid. By letting a current flowing in it a part of energy is stored. This operation of course cannot be too strong, in the sense that the current in play must be considerably small in order to avoid that the magnetic force created in the solenoid is not able to move the armature of the actuator causing a continuous feedback on the touch surface. The simplest way to create a charging circuit and drive the solenoid is by using a low side n-channel mosfet connected directly to ground. The nmos is driven by a mcu and connected to it though an RC network as it is possible to see in the picture below.

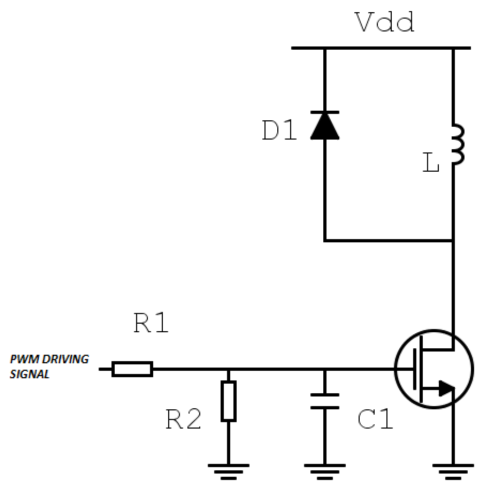


Figure 3.2. Charging circuit

R1 is used to avoid switching loss due to the fact that the mosfet does not turn off quickly since there are some parasitic oscillations caused by gate capacitance and the inductance of the connection. The value of this resistor can be also small in the order of hundreds of Ohms. R2 instead is commonly used to reduce the gate -source voltage to 0 V when the driving signal is open circuit. The Capacitor C1 is connected between the gate and ground to improve the turning on of the nmos.(da rivedere) Vdd is the supply voltage that is 12 V, a value that is normally use as external voltage and that can easily find in the vehicle cockpit.

This part of the circuit is the same used for the haptic feedback that is obtained with a pwm that drives according to the pwm value the current and so the force exercised by the actuator that is represented by the inductor in the circuit.
The study of this particular solenoid model is probably the most tricky challenge since it is composed not only by an inductor such in the ideal model, but also by a resistor and a capacitor in parallel. However since the capacitance is some orders of magnitude lower it can be neglected since the working frequency of the circuit is far from the resonating one. The series resistance has been estimated to be (5.4 ±0.01) Ω. This value is very important since it determine the current behaviour when the nmos is shorted to ground and so when the pwm signal is high.
From the Faraday' s law of induction it is possible to describe the phenomena according to which a current flowing into an inductor generates a voltage across it :

$$V_L = -\frac{d\Phi_B}{dt}$$

where the magnetic flux as previously explained is equal to $\Phi_B = L\ i$. According to this the equation can be rewritten as:

$$V_L = -\frac{d(L\ i)}{dt} = -L\ \frac{di_L}{dt}$$

This means that the voltage generated on the inductor is directly proportional to the rate of change of the current but also to the inductance value. For simplicity its change are not considered in the charging phase and so the derivative term $\frac{dL}{dt}$ is setted to zero. This approximation is correct if we consider the circuit in two different moments t1 and t2 in wich the value of the inductance is different but stable.
The circuit in this phase became the inductor connected to Vo from one side and the other one connected to ground through two resistances, the inductance one R_L and the nmos on resistance R_{ON}. According to the circuit depicted the following equation can be written:

$$Vdd = V_L\ +\ i_L(R_L\ +\ R_{ON})$$

$$Vdd = -L\ \frac{di_L}{dt}\ +\ i_L(R_L\ +\ R_{ON})$$

This is a non homogeneous first order differential equation that can be simply solved working in the s domain. The solution is :

$$i_L = I_0\ (1 - e^{-\frac{t}{\tau_1}})$$

$$V_L = V_0\ e^{-\frac{t}{\tau_1}}$$

The current behaviour increments exponentially to the steady state current value I_0 which is the obtained dividing the supply voltage by the sum of the resistances. In this case its value is almost 2 A. The time constant τ_1 instead is obtained by dividing the inductance by the sum of the resistors. The parameter that changes in the current equation due to the force pressure on the actuator is the time constant since it is dependent from the inductance value.

$$\Delta L \rightarrow \Delta\tau \rightarrow\ \Delta i_L$$

To quantify how much the inductance variation affects the current it is possible to study the following equation:

$$\Delta\, i_L = i_{L|L2} \;-\; i_{L|L1} = \; I_0\; (e^{-\frac{t\;R}{L_2}} \;- e^{-\frac{t\;R}{L_1}})$$

Supposing a variation of 3 mH in the value of the inductance, which normally is 1.85 mH, the relative change in the current maximum value at the end of the charging phase is about 20 mA. The first graph represent the different current profiles for the different values of

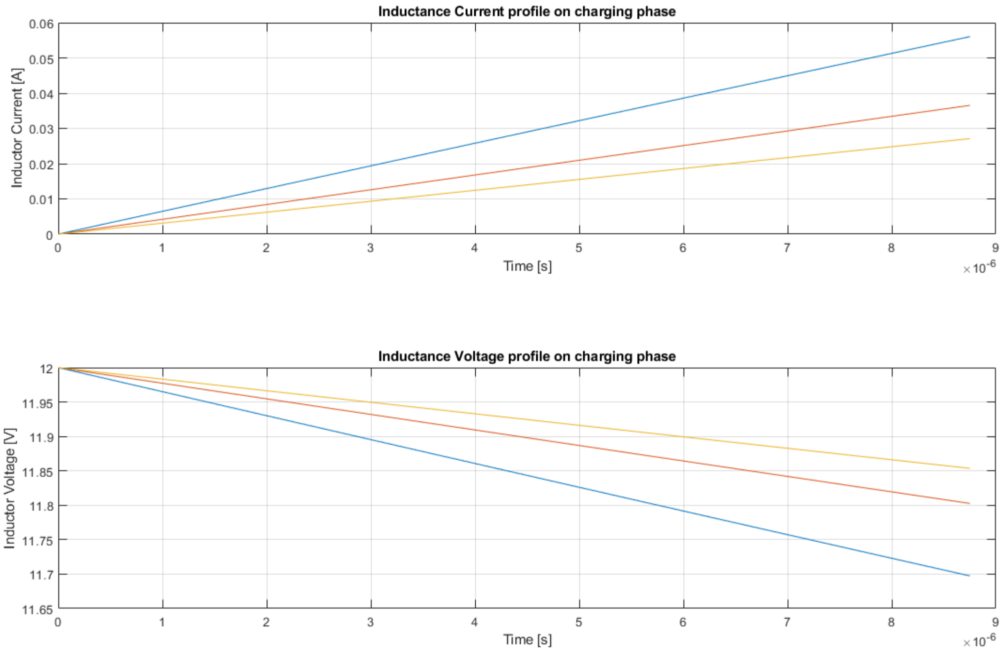


Figure 3.3. Different current and voltage profiles for different values of the inductance in the charging phase

the inductance and in particular the blu one is the current profile for L equal to 1.85 mH, the red 2.85 mH and the yellow one 3.85 mH. After 8.75 μs the charging phase elapse and all the energy stored in the inductor is discharged in a second circuit in order that can be measured.

3.2.2 Discharging and measuring phase

The aim of this part of circuit as previously anticipated is to measure the change in current/voltage manipulated by the inductor containing the information about the force applied. The first attempts were about discharging the current into a resistor and then using a differential amplifier read the voltage difference on it, compute the current and finally estimate the time constant and so the inductance. However this particular solution was not precise and so it has not been pursued. To obtain a better precision comes the idea to use a comparator that could be activated only above and below particular thresholds to convert the information into a pulse.Starting from this idea the following circuit has been developed.

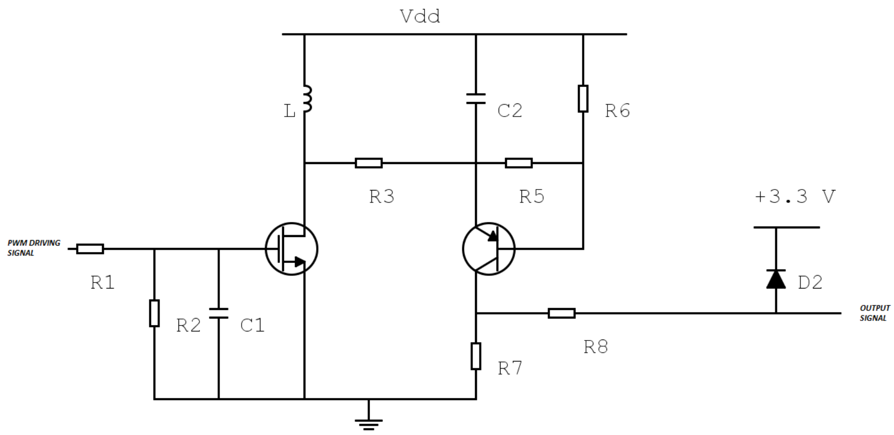


Figure 3.4. Measuring circuit

It is composed by a resistive network and a capacitor connected to the supply voltage whose aim is to create a voltage divider to control and manage the pnp transistor and activated it. The pnp when activated let the current flows from the emitter to the collector and then through the collector resistor causing a voltage drop on it. This voltage is then divided in the R8 resistor and to the diode.
Since there are three resistors connected between Vdd and the drain of the nmos transistor during the charging and the actuating phase a current flows in it and for this reason their value should be high enough to avoid an high power consumption[1]. Normally the value of the resistor R3 is used to control the voltage on the emitter and so R5 and R6 are tuned to reduce current flowing. R5 is also important since it provide the reverse emitter-base voltage when the transistor is in interdiction state. Of course also in the capacitor an impulsive current is present in the first phase however is very small and in series with R3 and R_{ON}. The first step is the study of the discharging current through the capacitor and the resistors. In fact the value of the capacitor should be small to avoid that circuit

[1]To avoid power consuption it is possible to replace R3 with a diode, with the anode connected to the nmos drain. In this way when the nmos is ”on” the diode is reversely polarised and no current flow through it.

resonates but on the contrary that it has an overdamped response. To better study the behaviour during the designing phase LTSpice simulator has been used.

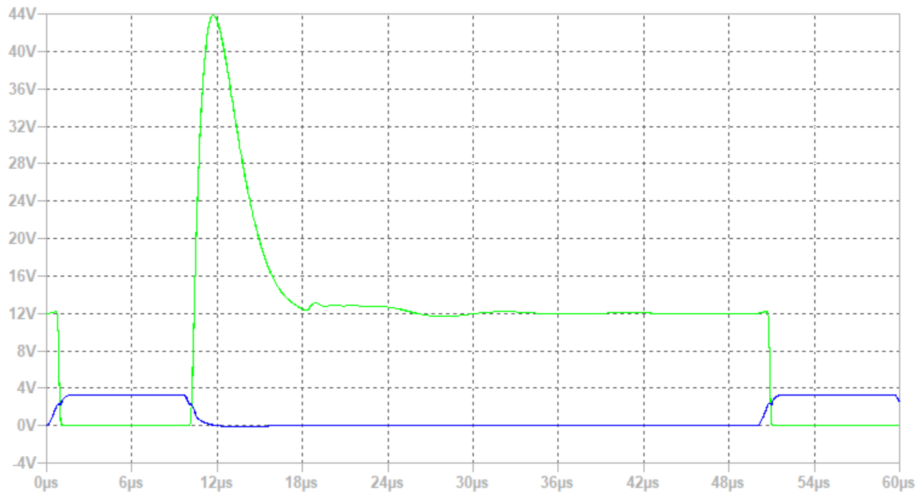


Figure 3.5. Simulated overdamped voltage response and driving signal

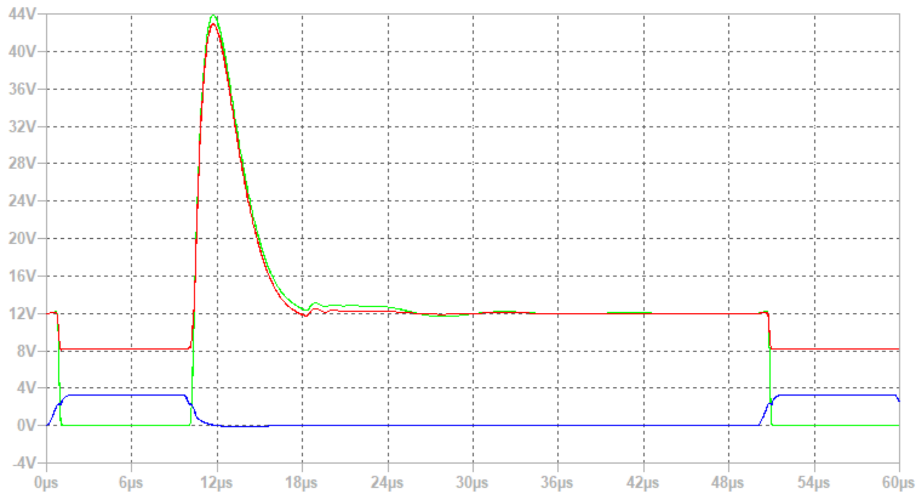


Figure 3.6. Simulated emitter and base voltage profiles

When the discharging phase starts the nmos is opened and the inductor due to its physical properties tries to maintain the current constant. According to the its constitutive equation the voltage rise impulsively in the same direction of the current, due to the minus sign present in the equation. The inductor so can be represented in this phase as a voltage generator in series with the supply voltage. Figure 3.5 represent the voltage on the emitter that for the reason just exposed is higher than the supply voltage. Figure 3.6 represents the emitter voltage and the base voltage in red. As it is possible to see during the charging

phase and so when the signal driving the nmos is high the V_{EB} of the pnp is negative and so the transistor is in interdiction mode, thus means that no current flows in it. On the contrary in the second phase the emitter voltage rise quickly than the base one and when the V_{EB} rise (up to 0.6 V) the transistor turn on.
Since the V_{EB} is the same voltage present on the R5 resistor (kΩ order) the current passing in it can be neglected and so it is possible to consider all the current entering in the emitter. The pnp however saturates because of the large quantities of current that it must handle.

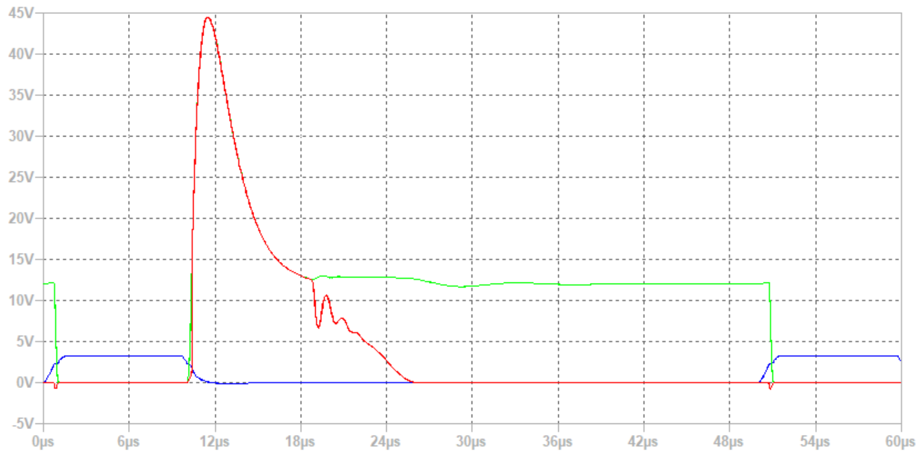


Figure 3.7. Simulated emitter and collector voltage profiles

This can be see in the previous figure where the red graph is the profile of the collector voltage and the V_{EC} is equal to 200 mV till the point where the pnp enter in the active region and so the V_{EC} increase. The transistor goes into the active region when V_{CB} decrease and became negative, while V_{EB} continue to be positive. The two emitter and base resistors and in particular the collector R7 allows to manage the transistor behaviour in such a way that it is more sensitive to the reluctance variations. In fact through the resistor it is possible to modify the transistor behaviour by deciding where activate it and where disable it, by entering in the cut off region.
Figure 3.8 shows a simulation about the effects of inductance variation on the emitter voltage and so also on the collector and base voltages. When the inductance increase, what is changing in the circuit is the time constant τ_2. Varying the time constant varies also the discharging current and that will be faster with a small τ_2 as show the blue curve in the graph which is the emitter voltage of the circuit with the higher value of inductance and so with the armature not pressed. Changing the time constant and so the current profile entering in the emitter of the transistor change also the collector current and voltage.

When the collector voltage becames higher than 3.3 V plus the V_D of the diode that is 0.7 V the current start flowing in it and so the output signal connected to the micro controller rise to 4 V. When at the contrary the collector voltage decrease under the same threshold

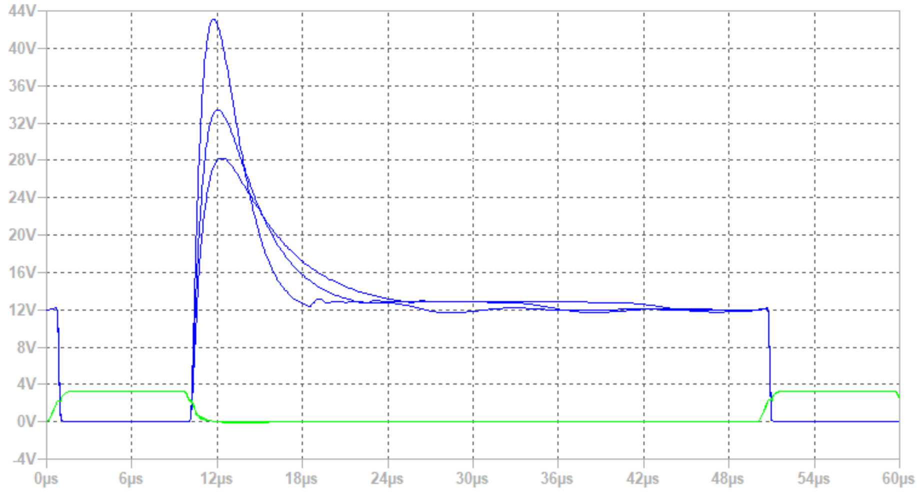


Figure 3.8. Effect of inductance variation on the emitter voltage profile

the diode shut down and the output goes to zero following the collector voltage slope.

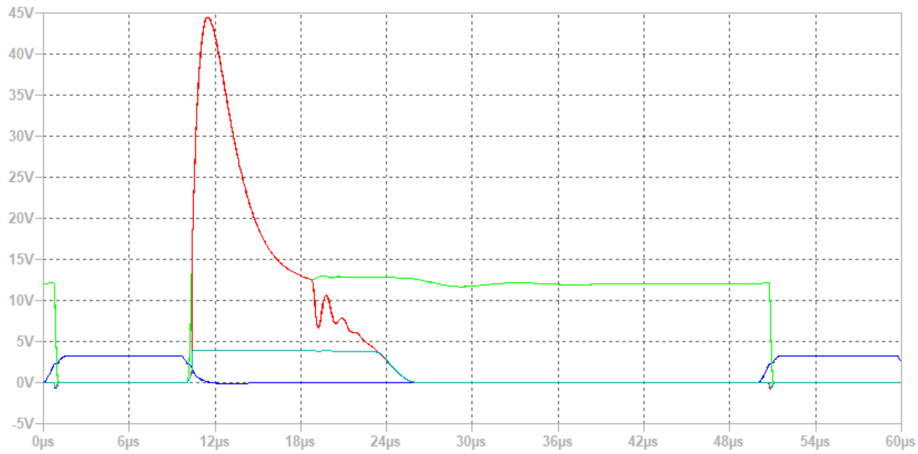


Figure 3.9. Output pulse profile

This last graph represent the final pulse generated that will be measured by the micro controller through an input capture procedure.

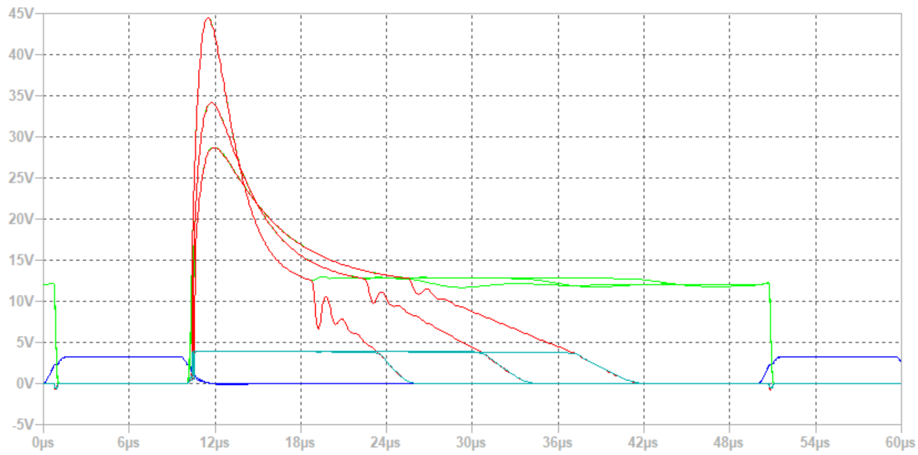


Figure 3.10. Output pulse profile with inductance variation

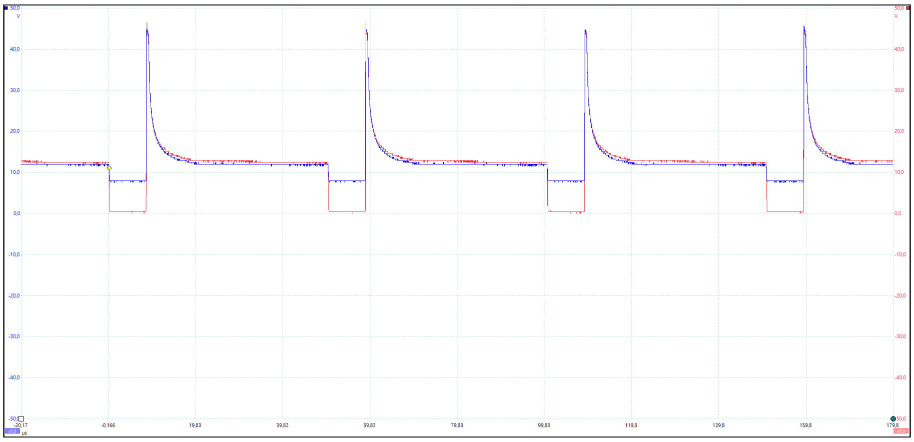


Figure 3.11. Emitter and base voltage profiles recorded with oscilloscope

3.2.3 Flyback diode circuit

In all the previous schematic an important part has not been included and this is the flyback diode. This diode is normally used in parallel to an inductive load to avoid the flyback effect that is nothing else than the back EFM voltage generated by the inductor when the current rapidly decrease/increase as described by the constitutive equation of the inductor $V_L = -L\,\frac{di_L}{dt}$. When this happens the diode became forward-biased and so start conducting and allow the current to recirculate avoiding that the voltage on the switch, in this case the nmos, rise too much since it can cause serious damages on the device or brake it.

However if this diode is present all the current will flow in it and so the measuring circuit will not be able to receive any information. For this reason come the need to revise this diode with a more complex circuit that allow to use the diode only when it is not necessary

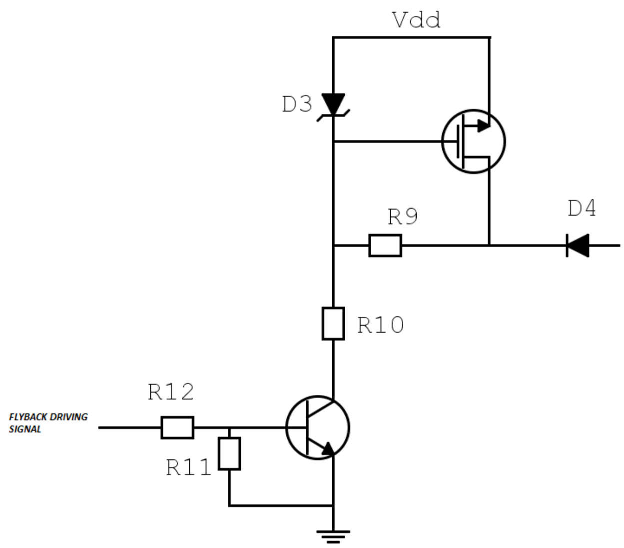


Figure 3.12. Diode activation/deactivation circuit

to measure and when the actuator must provide an haptic feedback. For this reason the circuit depicted in figure 3.10 has been developed.
The npn transistor has a resistor network at its input in order that can be driven by a micro controller by simply managing a voltage signal. If the driving signal is lower than a certain threshold that can polarize the base-emitter junction the npn is off and so no current will flow in the collector resistance. In this situation when the inductor voltage V_L rise above a certain threshold defined by the voltages on the two diodes and those on the resistance R9.

$$V_L > Vdd \ + \ V_D \ + V_{R10} \ + \ V_{thn}$$

When this happens the first diode D4 is forward-biased while the Schottky diode is reverse-biased meaning that the gate source voltage V_{GS} is positive and if higher than V_{thn} that the mosfet is in the on state and that a connection between v_{dd} and the diode is established such as in the normal flyback circuit schematic. This force the inductance current to flow in this part of the circuit avoiding that the driving nmos and the measuring circuit can be damaged. The second situation is when the input signal set to an high logic level and this means that the voltage on the resistor R11 is higher than 0.7 V and so the transistor turn on. The Schottky diode became forward biased fixing a negative V_{GS} that put the nmos in the interdiction zone. However a small current flow though the flyback diode and then through the two resistors to ground.
This means that a part of the current will not enter in the measuring circuit and this current can be controlled with the value of the R10 that can be grow to decrease this current.

Simulation in Fig 3.14 show in blue the pwm driving signal that charge and discharge the solenoid, in red the V_{GS} and in green the current flowing in the flyback diode. The following

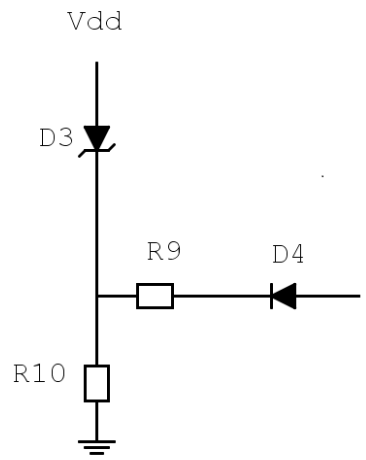


Figure 3.13. Flyback disabled circuit

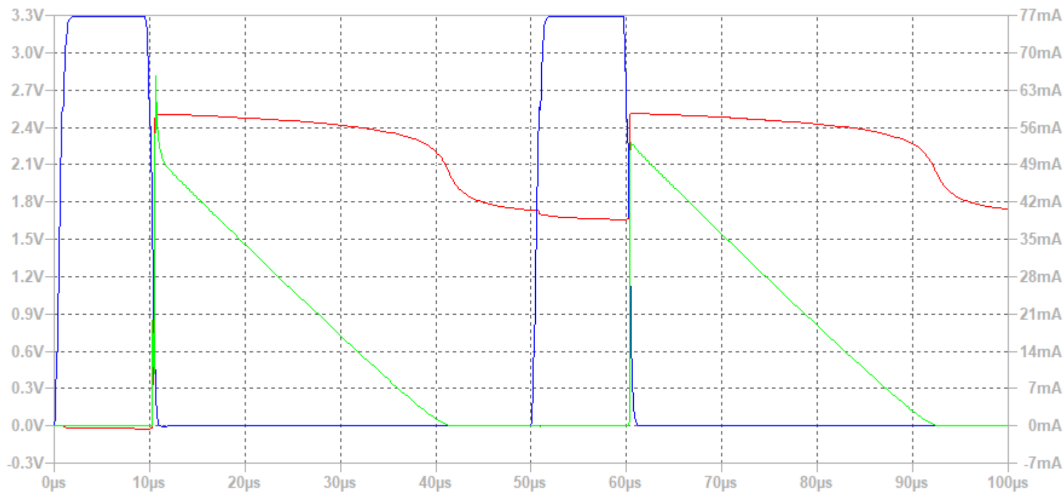


Figure 3.14. Current in the flyback diode and V_{GS} with circuit on

picture has been obtained simulating the whole circuit and switching the flyback circuit. In the first 350 μs the flyback driving signal (green one) is high and so the measuring circuit is working and needs to receive the information contained in the current. In the second part instead the flyback circuit turns on and so only a small part of the current flows and so of the information.

Of course this behaviour can be changed in a way that when the diode is activated is working no current flows in the measuring circuit, however this can be suitable. In fact this allows the measuring circuit to work even during the haptic feedback actuation. This allows to record all the data of the pulse and see if there is a degradation of the pulse, so if the energy transmitted though the pulse is lower. In fact since the data read are

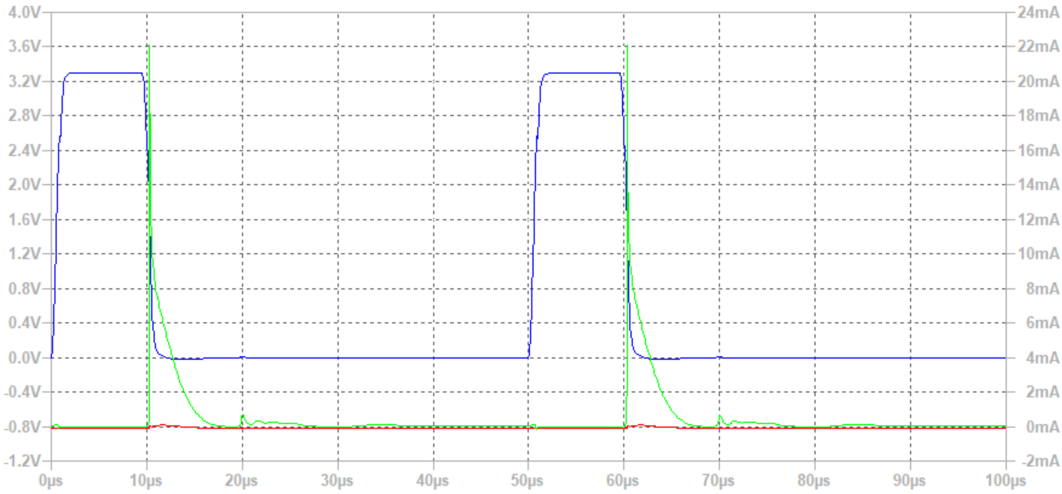


Figure 3.15. Current in the flyback diode and V_{GS} with circuit off

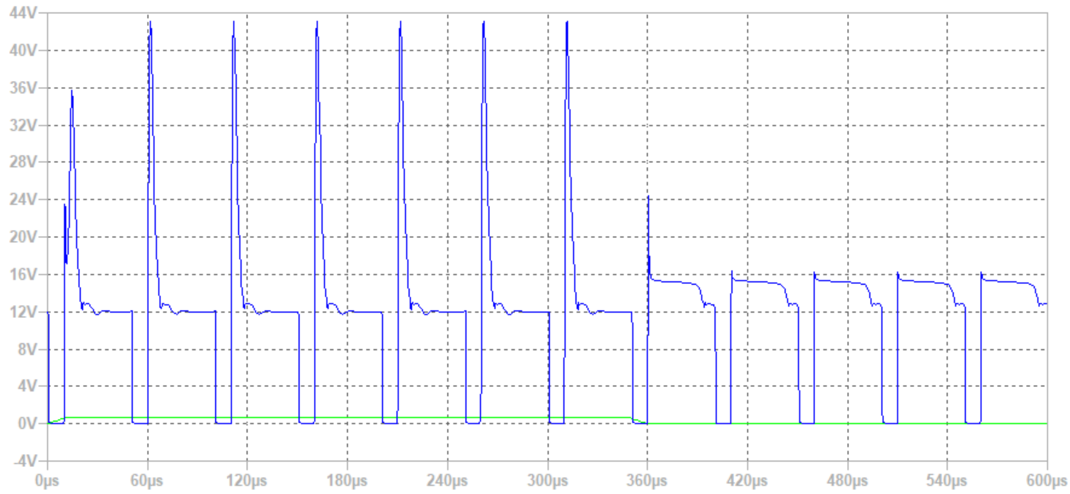


Figure 3.16. Voltage on the input node of the measuring circuit with a transition of the flyback circuit

directly proportional to the displacement, they are also proportional to the acceleration of the touch surface. Using an FFT on the device is so possible to understand if the frequency of the pulse spectrum are the same or are changed due to an accident or a mechanical defect.

3.2.4 Final system realisation

The final step is the merge of the two circuit, and the choice of the microcontroller. Each measuring stage needs 3 pins, two output pins for the control of the flyback diode and the excitation of the solenoid and one input pin for the reading of the pulses. The microcontroller chosen is the cypress CY8C4247AZI-L485. It is a 32 bit microcontroller with a clock frequency of 48 MHz, 256 kB of flash memory and 32kB of RAM.

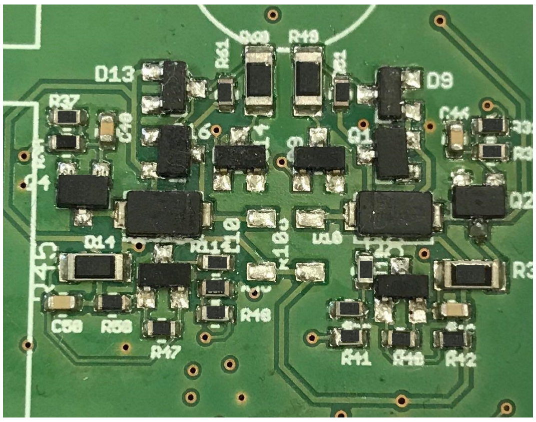


Figure 3.17. Pcb of two symmetric measuring stages

An important part is the communication system that is in charge of providing information outside the board. This is done through a CAN-bus communication, that is a standard in automotive field due to its strength to electromagnetic disturbances. CANbus is connected to a CAN-USB converter that allow to read data transmitted in the analyse them.

3.3 Firmware/Software Design

To create a precise and repeatable measuring system a great relevance has been grant to the firmware development that is the core of the system because it allows the possibility to recognize the human force applied to a surface. Firmware must be able to monitor all the different inputs from the circuit and to send signals whenever it is necessary, and for this reason its role is very important.
The management of timing, that is one of the most important aspect in a RTS (Real Time System) because it assure that all the operations of microcontroller and consequently of the measuring stage are performed at the right time. This management is possible thanks to the use of different Timer Events that are evoked when a certain amount of time has elapsed. In order to perform that an internal timer has been used. The prescaler for this particular timer has been setted to 1 and so the working frequency it is the same of the clock 48 Mhz with a correspondant period of 20,83 ns; at the same time an interrupt has been setted on a compare match in order that when the counter of the timer becames equal to a certain compare the interrupt is setted and the Event is evoked. In this case three different events are created:

- **1 ms Event** : this event is used as a trigger for the Measuring Position task that compute and store the information about the pulse length, for the baseline task used in the Force recognition algorithm
- **5 ms Event** : this event is the trigger for the haptic feedback task that takes care about the management of the different output feedbacks
- **10 ms Event** : used for temperature measurement tasks or others tasks which not require a frequent call

This timing management allows to synchronizes each task and be sure the order of execution is always the same.
Another important solution adopted is the structure of the generic task that is based on a Finite State Machine normally adopted in Hardware design but that can be a suitable solution for firmware applications like this where different task needs to talk each other. In this way it is more easier to understand which is the actual state of each task and so it easier to share data between them. Each task has normally 5 different states: This solution allow to avoid thousands of "if-else" statement in the code and let the code be easier to read, maintain and test. A task model code adopted is reported below:

```
void Task(void)
{
        switch(Task.sts)
        {
                case TASK_STS_INIT:

                /** Init functions**/
                if()
                {
                        Task.sts = TASK_STS_RUN;
                }

                break;

                case TASK_STS_RUN:
```

```
                /** Run functions**/
                if()
                {
                        Task.sts = TASK_STS_WAIT_SLEEP;
                }

                break;

                case TASK_STS_WAIT_SLEEP:

                        Task.sts = TASK_STS_SLEEP;

                break;

                case TASK_STS_SLEEP:

                /** Sleep functions**/
                if ()
                {
                        Task.sts = TASK_STS_INIT;
                }
                break;

                case TASK_STS_ERROR:
                /**Task in safe mode*/
                break;

                default:
                break;
        }

}
```

Since the function "Task" is called in a timer event, each time the event is evoked the state of the FSM is updated according to some parameters expressed in the if statement. In this way in every moment, knowing the inputs and the timer output, it is possible to predict the state of each task. This is very useful during debugging phase where it is possible to monitor the state signal understanding which condition is true or false. Firmware has been structured so that each functional group is divided in a particular task, allowing watertight compartments that in case of error does not shut down the whole code. In the next pages the most important tasks would be analysed.

3.3.1 Measuring Task

The first task is in charge of the managing the measuring system in order to obtain the information about the position of the metallic pad. Managing the measuring systems do not means only read the informations but also provide the correct signals to the circuit in order that it can correctly work. The first operation performed in fact is to turn off the clamping diode circuit by writing a logic "1" on the signal that is connected to each of the four nmos. In this way the circuit is prepared to the measuring phase because all the current, and so the information, contained in the inductor is no more pushed into the diode but can liberally flow in the measuring circuit. The second step is to properly set the signal that drives the other four nmos transistors that control the current flowing in the inductor and that are the same used for the creation of the haptic feedback.
The signal that drives each nmos is created with the use of Pulse With Modulation in such a way it is possible to change the duty cycle values quickly, that is required to let

unchanged the possibility of creating different pulses for different situations. In the initial state so the pwm module is instantiated, the frequency is setted to 20 kHz and the duty cycle to 17 % (correspondent to 8.75 μs) and finally the pwm is enabled. Then an input capture timer has been setted in order to measure the pulse width of each pulse. The

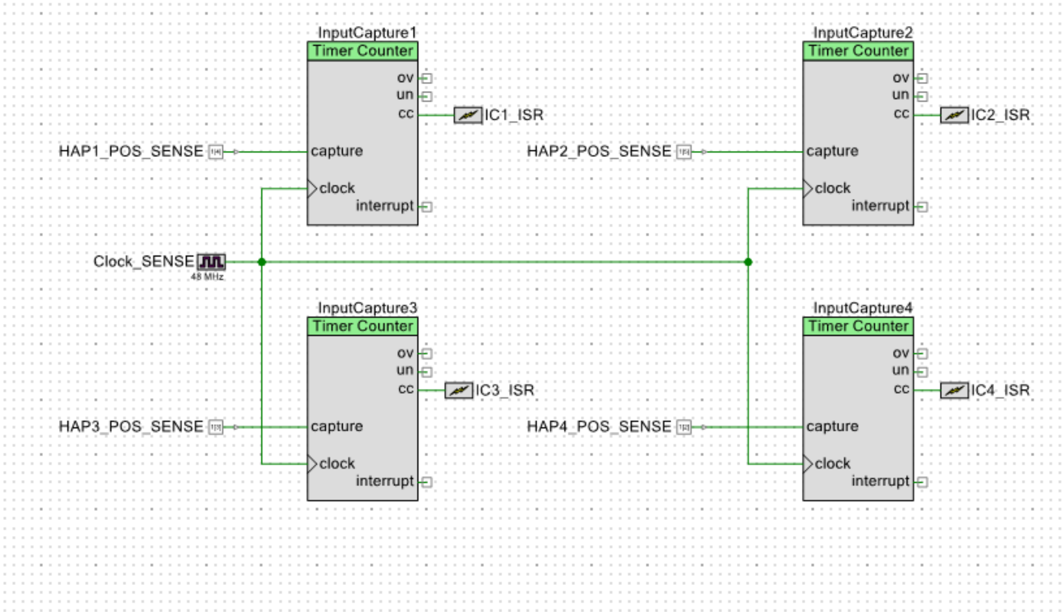


Figure 3.18. Input capture diagram

capture has been setted up without a clock prescaler (x1 prescaler) in order that it can works with the maximum frequency 48 MHz and so the resolution is the highest possible. The capture has been setted to trigger on rising edge of the input signal and when this happens an Interrupt is evoked. Finally both the pwm, the input capture and the ISR Interrupt Service Routine are enabled. This is the last operation performed in the initial state of fsm.

At the next clock cycle the fsm will update his state (running state) and so the measures can start. The measure itself is performed by the ISR that each 50 μs compute a new data. Since the first time the trigger is set to sense the rising edge, at the first rising of the signal HAP_POS_SENSE the Interrupt routine check if the edge is the correct one and if correct it store the value of the counter in a variable and set the falling edge as trigger. In this way when the signal goes down the routine is called back and compute the difference between the new value of the counter and the old one previously stored.

```
ISR(IC1_ISR_Interrupt)
{
        if((InputCapture1_TRIG_CONTROL1_REG & 0x0003) == 0x0000)
        {
                RisingValue1 = InputCapture1_ReadCapture();
                InputCapture1_TRIG_CONTROL1_REG |= 0x0001;
        }
        else
        {
                if((InputCapture1_TRIG_CONTROL1_REG & 0x0003)== 0x0001)
```

```
                {
                        DeltaPulse1 = InputCapture1_ReadCapture()-RisingValue1;
                        DeltaPulse1 &= InputCapture1_16BIT_MASK;
                        if(DeltaPulse1<1980)
                        {
                                WriteQueue1(DeltaPulse1);
                        }
                        InputCapture1_TRIG_CONTROL1_REG &=0xfffc;
                        }
        }

}
```

A control to prevent false signals is used, in fact the data is stored only if its value is lower than the maximum admitted value that is obtained as $f_{CLK} * (50\mu s \; - \; 8.75\mu s)$ that is 1980. In automotive field it is very important to prevent any type of accident and so lots of controls are inserted in the algorithm. If any error occurs the data can be stored in a ring buffer. This circular buffer is a data structure based on a FIFO and with a fixed maximum size. It is composed of two pointers, one for the write and another for the read that enable the program to write all the data and then read them until the two pointers points the same memory cell.
Since the pwm frequency and so the frequency of this measurement system is set to 20 kHz, this means that every 1 ms the microcontroller receives 20 data from each stage circuit and so each circular queue must have at least this number of storing locations. So each millisecond, the period of this task, every queue is read in this way. The program enters in a for statement that read data from queue and sum them until it reach a parameter called dynamic gain (in the tests fixed to 8) and then provide the sum as input to a moving average filter which is considered by the program the Raw value.
The moving average filter also called "box car filter" is a FIR Filter with a fixed coefficient. The formal expression is given by :

$$y[k] = \sum_{n=0}^{N-1} h[n]\ x[k-n]$$

$$h[n] = \frac{1}{N}\ ,\ n\ =\ 0,\ 1, ..., N-1$$

N is the filter lenght.
This filter is really common in DSP domain since it has a simple and weight implementation and it is optimum for reducing random white noise maintaining the sharpest step response. This last aspect is very important since the aim of this filter is to reduce the random noise that affect this type of measure but it must assure a precise fit of the input signal to not waist information about the system behaviour.
Since this is a low pass filter it is possible to compute the cut off frequency:

$$f_{cutoff} = \frac{f_{sampling}}{2\ \pi N}$$

Since $f_{sampling}$ is equal to 20 kHz, to fix a cut off of 50 Hz the number of samples N is set to 64.

3.3.2 Baseline task

While the measurement task provides all the informations and data from the sensing circuit without doing any manipulation of them, the baseline task is the core of the force recognition algorithm. In fact it receives as input from the previous task each 10 ms the Raw value and produces as output two other signals that would be later used by the Force task. These are the Tracking signal and the Baseline signal. The tracking signal is obtained by filtering the Raw signal with another Moving average filter placed in order to create a signal wich is slower with respect to the first one. In this case the number of samples is set to 100 in order to cut over 30 Hz. In practice this signal is slower but faster enough to see a rapid touch pulse.
However this signal is only used to manage and as control signal for the baseline.

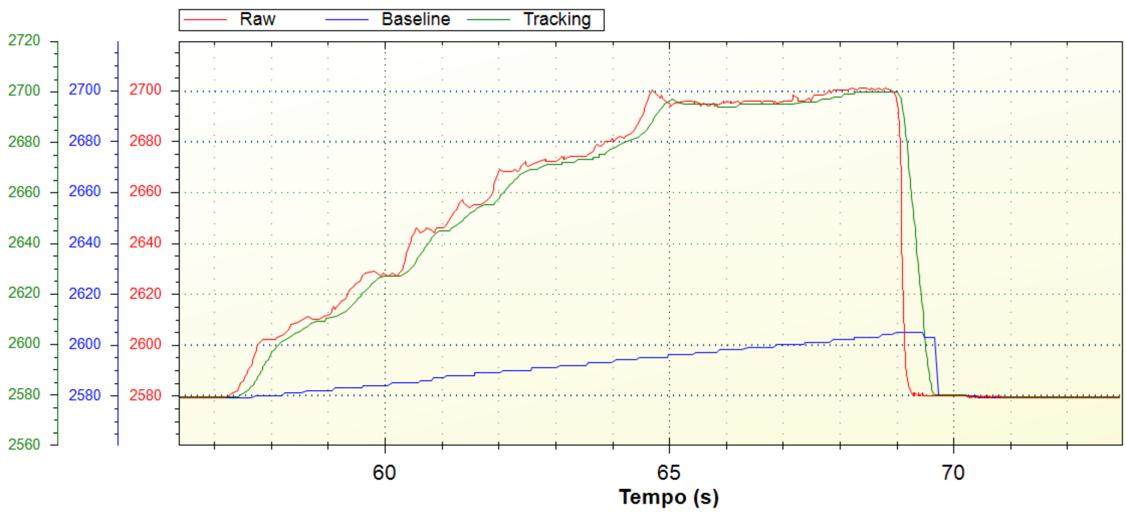


Figure 3.19. Raw, Tracking, Baseline signals when a force is applied to the sensor

The Baseline is a signal that help the microcontroller in the computation of the applied force. Why there is the need of a baseline?
This need is due to the fact that the raw signal can have a drift due to the temperature and hysteresis dependency and so for this reason it is not possible to simply see the Raw behaviour. In fact the signal may rise slowly due to a temperature drift without any force applied. In this case the system may interpret it as a slow force by comparing this value to a fixed value, but this is not correct.

So for this reason derives the need to have a moving threshold that follow the signals. In order to better design this signal many acquisitions are carried out to better understand the user touch pulse parameters, in order that the system can recognize them.

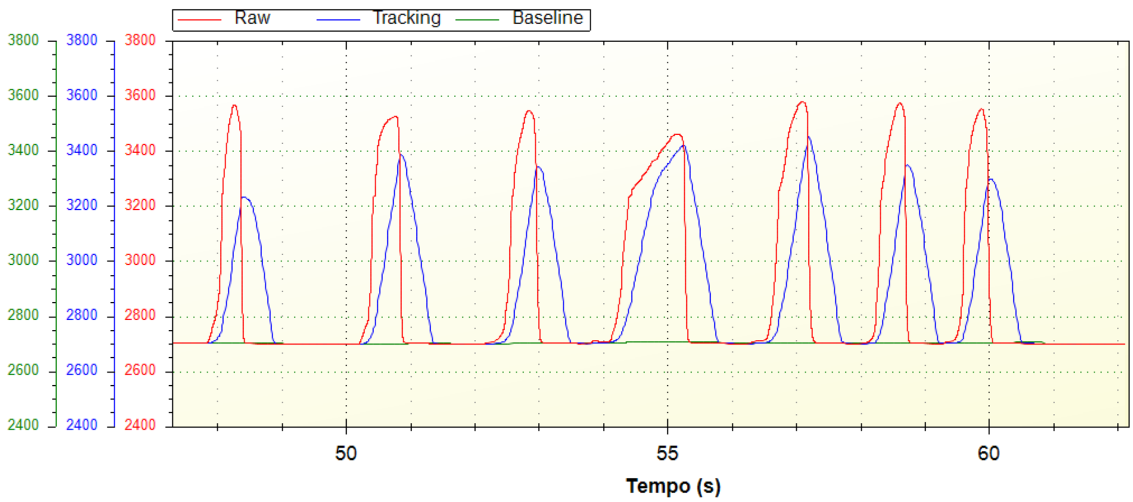


Figure 3.20. Different consecutive pulses recorded

Touch pulse

About that an acquisition campaign has been setted up to measure the user touch pulse, study its shape, velocity, pulse width and finally obtain an average touch pulse. This acquisition has been done using a moving structure on which there were placed an accelerometer and a laser measuring system to measure the displacement of the touch surface. Using the data acquired with the laser position instrument the following graph has been obtained.

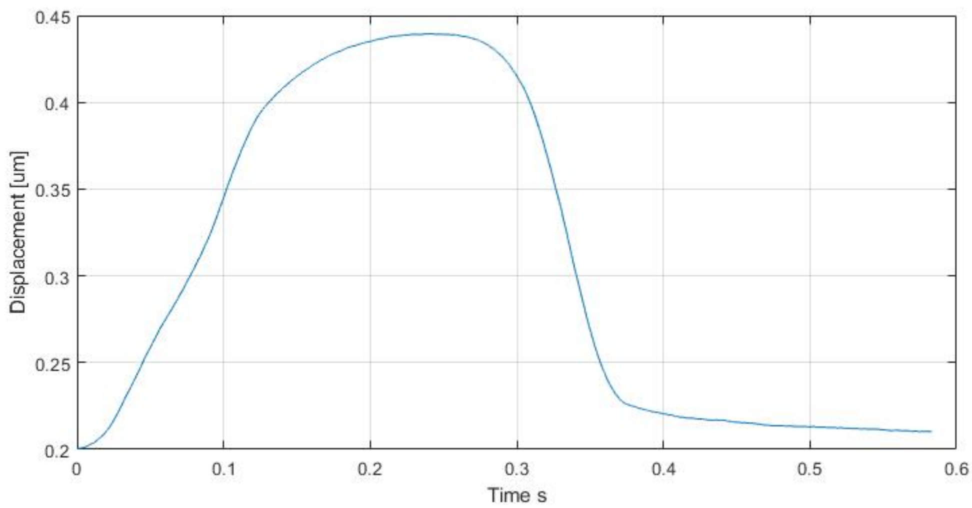


Figure 3.21. Finger pulse recorded with IPE acquisition system

This study allows to understand how slow the baseline should be, and in particular it must

have a cut off frequency of 2 Hz. So this means that each 500 ms it is updated in this way: the value of the baseline is set at the beginning equal to the raw value, then each 500 ms a procedure check if the raw value is higher than the Tracking value, if this condition is true the baseline is updated by incrementing its value by one. Another particular condition is when the raw value is lower than the Tracking value and the baseline.
In this particular case since we are not interested in the reading of traction forces the baseline is continuously updated in such a way that it is never higher than the raw value, but at least equal when no forces are applied.

3.3.3 Force task

The last important task is the force one, which is in charge of the computation of the force applied. This task receives as input the raw signal and the baseline signal and from their behaviour compute the final force. The force algorithm look at this two signal in search of an incrementing difference between the raw signal and the baseline.

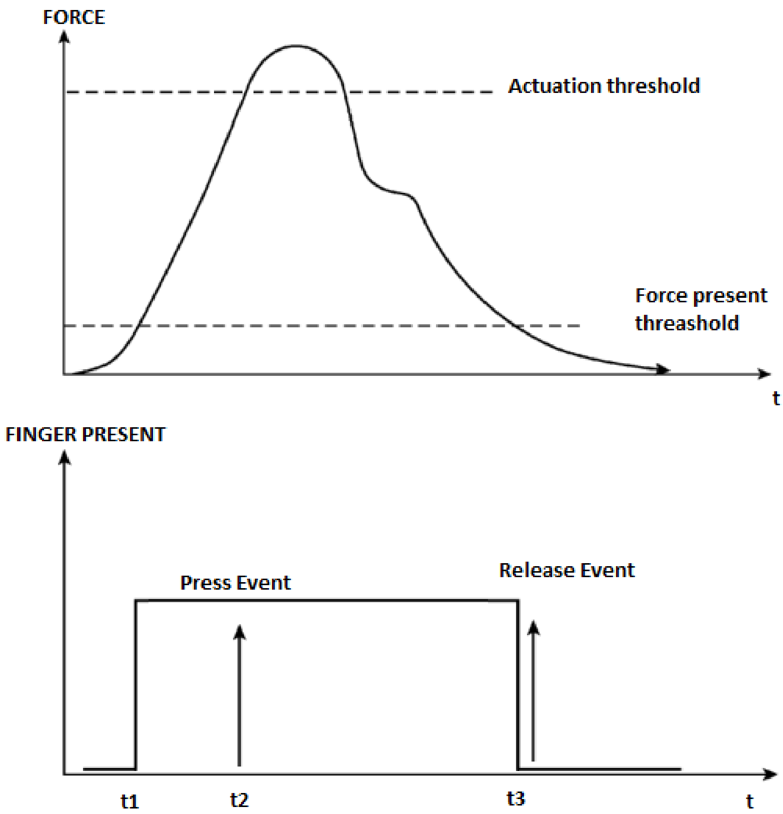


Figure 3.22. Force graph and relative thresholds

It determines that a force is present when the differential signal is higher than a certain threshold called "Force present threshold" that is normally a very small number. If this

threshold is overcame the signal of Force present is shared with the other tasks such the haptic one which starts the preliminary operations in order to be ready to provide the feedback when the a second threshold is overpassed. When the force signal arise above the second threshold, Actuation threshold, the processing algorithm generate a press event, that means that the force applied is sufficient to press the virtual key for example.Finally when the force signal decrease below the force present a Release event is generated which means that the user is releasing the finger from the touch. This configuration provide the possibility to have multiple press events while the finger present signal is high, however it can be changed in order that the only a press event can occurs and so a release event is needed to reactivate the button function. So the only parameter that the system needs to operate is the value of the actuation threshold. This can be done in production phase by a calibration step in which the desired force threshold is applied to the device and the value read by the device is stored in a non volatile memory, such the micro controller ROM in order that it becomes a fixed parameter.
Normally these thresholds are expressed in a dimensionless number since it is not necessary to express the force signal in Newton unless it is a specific application that requires it. In this way we avoid transform operations which are useless in this case. However a negative aspect is the fact that due to physical properties of the actuator each one can differ a little from the others and so these threshold are different from one stage to another. Since it is a very small difference if the required precision is not high it is possible to use a single constant for all the stages. In the opposite case each actuator and stage requires to be calibrated. This part will be analysed later in the testing chapter.

3.3.4 Other tasks

There are several other tasks running, starting from the one which is in charge of provide the haptic feedback to the one which takes care of the communication through a CAN-bus interface. Another task implemented is in charge of the measurement of the temperature through an ADC converter connected with a NTC resistor that is used to correct the samples that are affected by temperature drift.
Another interesting task, that in the future can be realised, could monitor the pulse produced during the haptic feedback procedure and verify its response. In this way the system can check if the energy released by the actuator is always the same or it is changed. This particular operation can be done since the measure provided by the circuit is directly proportional to the displacement and so implementing an FFT algorithm on board it is possible to see the energy of the signal and the frequency of work. In this way the system can perform a diagnostic in particular situation by itself.

Chapter 4

Testing and verification

After the design phase the whole circuit has been realized in order that a testing procedure can be applied to verify its behaviour and test it in various conditions. This is a very important part since it must check every possible situation the circuit could be subjected to. For this reason testing has been divided into multiple sections that cover all the required tests a new device should be submit in order to be released on the market. Since the starting field of this project is the automotive one, this phase is intended to verify the behaviour of the circuit in all the situation required by car makers specifications.

4.1 Repetitiveness Test

The first step performed has been the verification of the correct behaviour of the circuit.To perform a correct testing procedure a special structure has been built in order to be able to compare different results avoiding that random errors can occur. The first step performed is the test of the repetitiveness that is very important in a measuring system since if not repetitive cannot be used. For this reason different measurement campaign have been performed in different conditions (temperature, supply voltage). To have a more precise model avoiding errors due to the elastic constant k of the O-Ring the first measurements has been performed without it in such a way that the characterization is not bound to a specific spring.In this way the data provided by the system is only related to the distance between the metallic body and the metallic pad. This allow to use this data whatever type of spring is used. The decision to characterize the displacement and not the force has been taken with the aim to reduce systematic errors such the non-homogeneous application of the force. With the displacement instead it is enough using different non-metallic foils of various thickness. The normal operating range is about 100-600 μm and so tests have been executed in the range 0-800 μm with a 40 μm steps.
At first measurements with the same circuit stage and the same actuator at different time have been performed to understand if the system response is always the same or decay with the time. Fig. 4.1 show the results of 5 different characterizations performed in these conditions and as it is possible to see the trend is the same. Statistical analysis is then used to analyse each interval data and produce numerical informations.

The first operation is the computation of the standard deviation σ that provides an

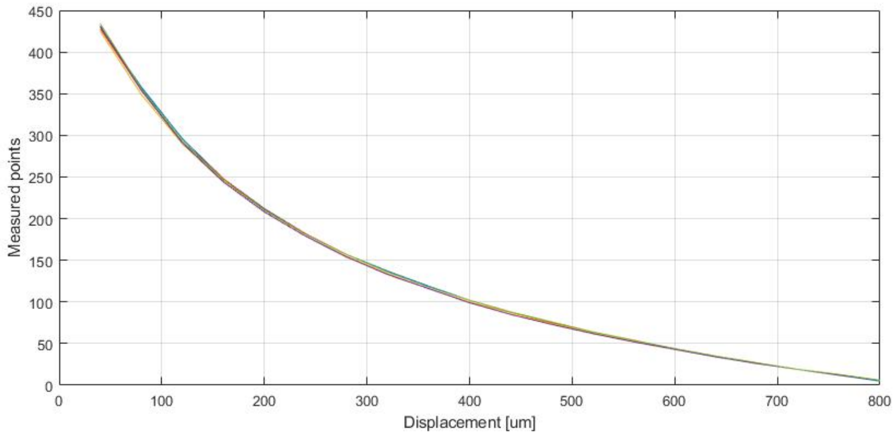


Figure 4.1. Characterization curves of the same NC actuator

information about the variation of a set of data and it is a good index of the random error in the measurement of a physic quantity.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}, \qquad where\ \mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

If its value is low it means that all the data tends to be close to the average μ, at the contrary when its value is high it means that there is a dispersion. In the table below the standard deviation is reported in the first row.

40 µm	80 µm	120 µm	160 µm	200 µm	240 µm	280 µm	320 µm	360 µm	400 µm	440 µm	480 µm	520 µm	560 µm	600 µm	640 µm	680 µm	720 µm	760 µm	800 µm
Standard deviation σ																			
3,361547	3,34664	2,302173	2	1,788854	1,516575	1,414214	1,643168	1,341641	1,414214	1,30384	1,140175	0,894427	0,707107	0,447214	0,547723	0,447214	0	0,547723	0
Average																			
429,6	354,2	292,6	246	209,8	180,6	156	135,2	117,4	101	87,2	74,6	63,4	53	43,8	34,6	26,8	19	12,6	6
Variation coefficient σ*																			
0,782483	0,944845	0,786799	0,813008	0,852647	0,839743	0,906547	1,215361	1,142795	1,400211	1,49523	1,528385	1,410768	1,334164	1,021036	1,583013	1,668707	0	4,347004	0

Figure 4.2. Analysis of measured data

However it is not an effective way to understand data. For this reason normally it is defined a variation coefficient σ^*.

$$\sigma^* = \frac{\sigma}{\mu}$$

This parameter, expressed in percent, express the precision and repeatability of measured data, and so of the measuring system. In fact as depicted in the table there is variation in the order of 1 point percent that means that with 1 N applied it has an incertitude of 0.01 N. The same procedure has been executed changing both the actuators and the measuring stage. This operation allow to understand the compatibility of the system and the fact that a common result is provided by different stages with the same input. Figure

4.4 represent the results of this acquisition campaign.

40 µm	80 µm	120 µm	160 µm	200 µm	240 µm	280 µm	320 µm	360 µm	400 µm	440 µm	480 µm	520 µm	560 µm	600 µm	640 µm	680 µm	720 µm	760 µm	800 µm
Standard deviation σ																			
20,25586	16,16168	13,0499	10,48809	8,700575	7,155418	6,041523	5,612486	4,636809	3,911521	3,577709	3,082207	2,774887	2,280351	2	1,516575	1,30384	1,30384	0,83666	0,547723
Average																			
403,6	337,2	280,6	237	202,8	174,8	151	131	114	98,4	84,6	72	61,2	51,2	42	33,6	25,8	18,2	11,8	5,4
Variation coefficient σ*																			
5,018797	4,792907	4,650714	4,425354	4,290224	4,093488	4,001009	4,284341	4,067377	3,975123	4,22897	4,280843	4,53413	4,45381	4,761905	4,513616	5,053645	7,163959	7,090339	10,14301

Figure 4.3. Analysis of measured data

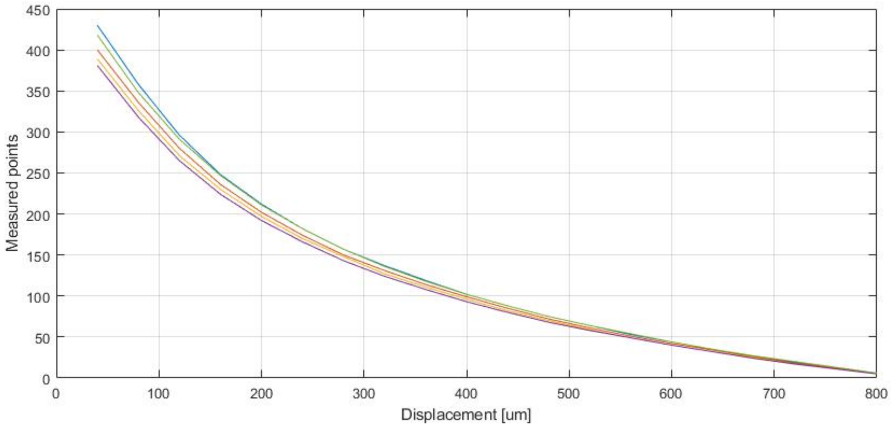


Figure 4.4. Characterization curves of different NC actuator

In this case the results are a little bit different and this can be seen in the σ^* coefficient values which are always lower than 10 % , that is a constraint of the system, except in the range higher than 760 μs. This effect can be explained by the fact that due to the production process limits that affect the internal parameters of the actuator. For example an error of the coil turning machine could affect both the internal resistance (negligible effect) and its inductance modifying so its response. However as it is possible to see from the graph the behaviour correspond to the same as before.
After this results it is possible to conclude that if the application required precision rate is at most 1 % the only solution is to characterize each measuring stage in production phase and saving the fitting parameters of the model in the ROM of the micro processor. For all the others applications since this device is used as a force threshold detector this is not needful. However this study has been fulfilled in order to verify that this measuring system can work also in others applications where an higher precision is required and of course repetitiveness.

4.2 Temperature Test

The following tests concern the temperature behaviour of the device. This is really important in an automotive application since the range at which the circuit should work correctly is between [-40:+125] °C. Of course correctly means with a variant coefficient whose value is lower than a specified threshold proper of the application.
To perform this test another structure that allow the heat to warm the actuator has been realized. Of course the better solution is a climatic chamber, but this test was not possible during the thesis time. However this solution allow to estimate the possible problems and the behaviour with the temperature changes. To heat the system a 50 W resistor has been fixed below the metallic structure in order that it can warm both the structure, the actuator both the air around the circuit board. A temperature NTC circuit has been built to estimate the temperature and compare its value with the thermal camera used to verify that each part is correctly subjected to the heating. This solution with a power resistor allow to better control the heat by simply modifying the power dissipated in it thanks to a variable DC power supply.

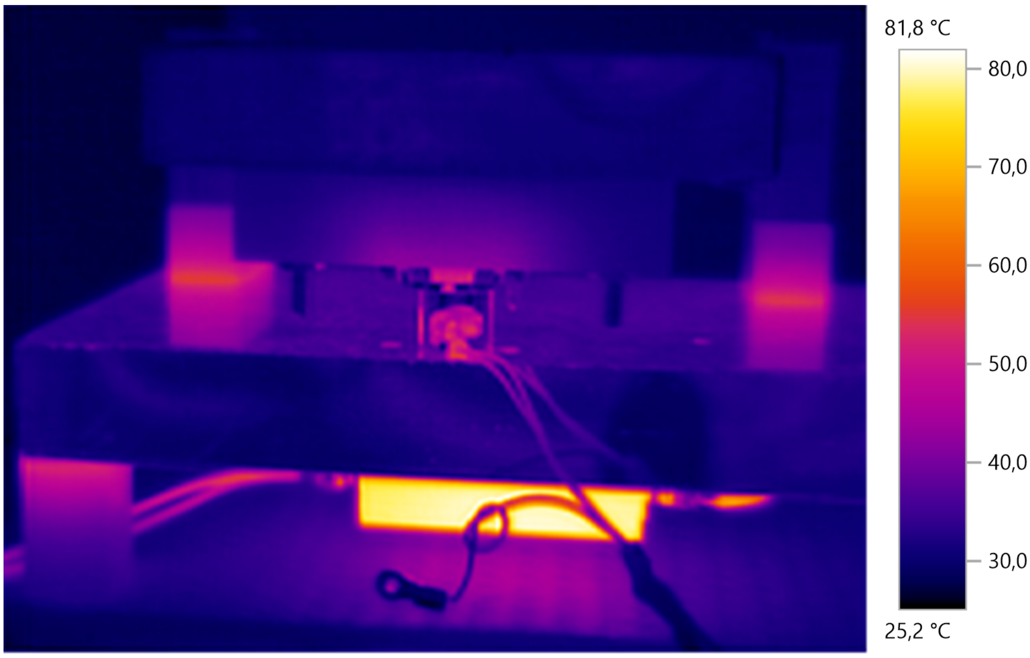


Figure 4.5. Setup of the temperature test acquired with a thermal camera

As in the previous measurement campaign a displacement measure is performed but in this case this procedure is repeated at each temperature interval. To speed up the process the displacement has been incremented to 80 μm and the steps reduced for simplicity. At the end the shift or difference with respect to the 25 °C respective data is computed. Finally a fitting is performed to build a model that can represent the temperature behaviour. Fig. 4.6 show the thermal drift and allow to see that with temperature rise the measured value also increases. This effect is more evident when the distance between the body and the pad is small and so represented with the blue curve.

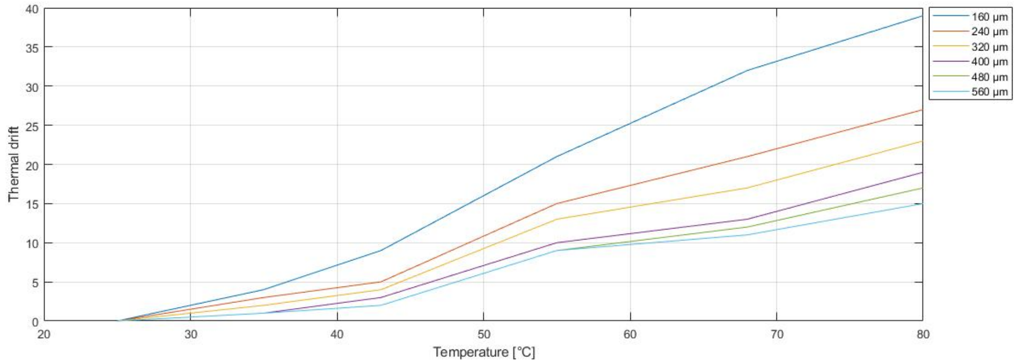


Figure 4.6. Relative increment for different displacements at different temperatures

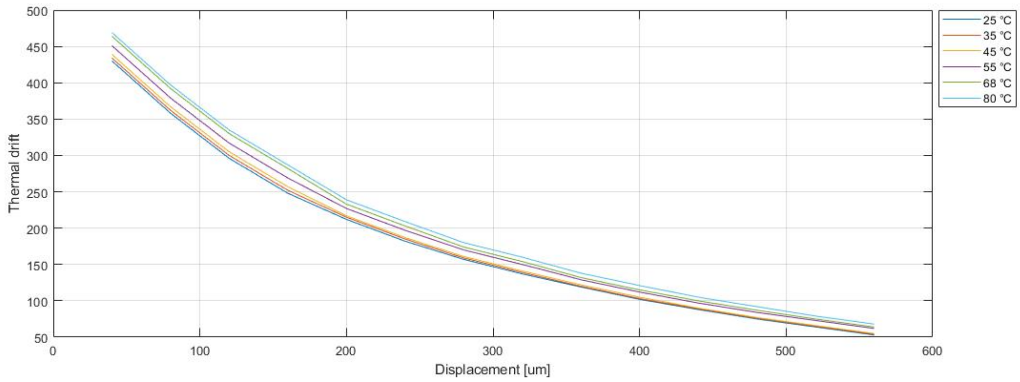


Figure 4.7. Thermal drift graphs

The second graph instead is the effect of the temperature change reported on the characteristic proper of the sensor. As it is possible to see the effect of the temperature can be simplified as a shift upward of the characteristic, even if for displacement below 160 μm there is also a change in the slope. However since the sensor return a differential measure the effect of the temperature on the final measure is not perceivable or can be considered negligible.

$$y_{measured} = (y_2 + \delta y_{1Temp}) - (y_1 + \delta y_{2Temp}) \approx \Delta y$$

Of course this is an approximation but since the variant coefficient σ^* and the relative error for displacement in the range [200:560] μm stay below 10% it can be considered acceptable. In other applications where the required precision is higher a possible solution is to monitor the temperature behaviour and introducing correcting factors that can be added during the computation of the measure. This can be done with an NTC resistor biased though a resistor and read with an ADC input of the microcontroller. The measured temperature will then point to a matrix in which for each temperature there are the correspondent corrector factors.

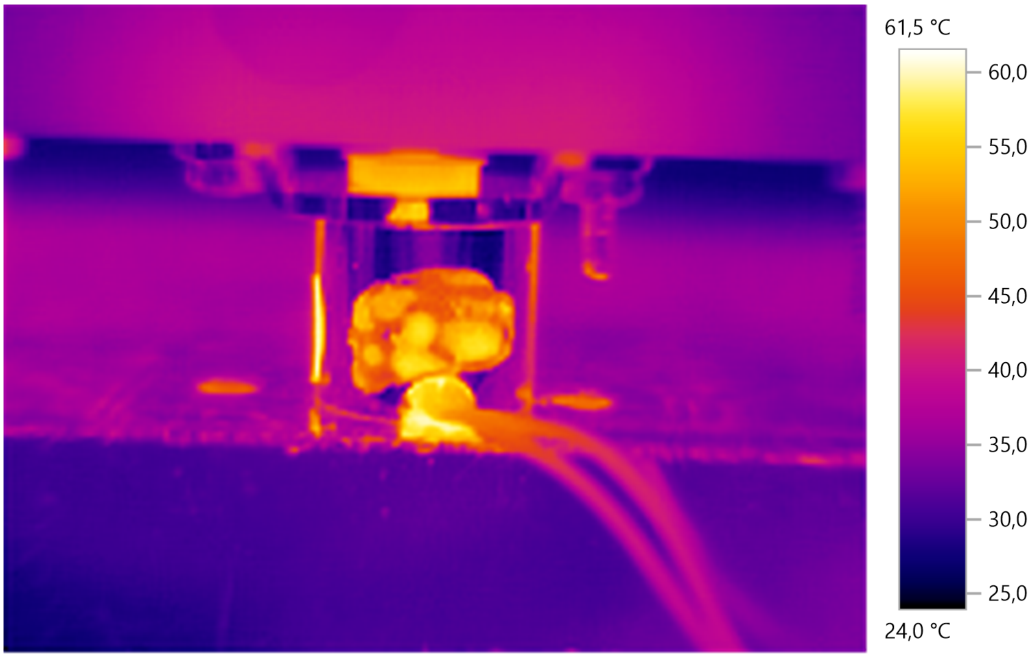


Figure 4.8. Temperature measurement with a thermal camera

Finally considering the force present signal generated by the system, the temperature has no effect on it since the baseline is able to track the raw signal that may increase due to temperature. Of course this is true in a plausible situation where the temperature has a maximum change rate of 1 °C/s.

4.3 EMC testing: Emissions and Immunity

The final step of the testing procedure is the Electromagnetic compatibility test, which is a required test for any commercial device. It is divided into two parts: Emission and Immunity/Susceptibility tests.
Emission testing is used to verify that the device under test is compliant to the standards and laws. It check if there is some type of electromagnetic emissions produced by the dut both radiated in the air, that can be measured through an antenna, both conducted through a cable to the power supply lines. These tests are performed to be sure that the device would not affect the behaviour of others devices working in the near or connected to the same power supply. Immunity testing instead is intended to verify the robustness of the dut, in order to be sure that it will continue to work correctly even when emitted radiations or disturbs on the supply lines are present.

4.3.1 Radiated Emissions

Radiated emissions are related to the unwanted, non-intentional emissions of electromagnetic energy of a device. These emissions could affect the behaviour of other devices in close proximity and for this reason these types of test are performed. To realize this type of test an anechoic or semi-anechoic chamber is required to be sure that all the emissions are received by the antenna and avoid to receive unwanted emissions from other devices.

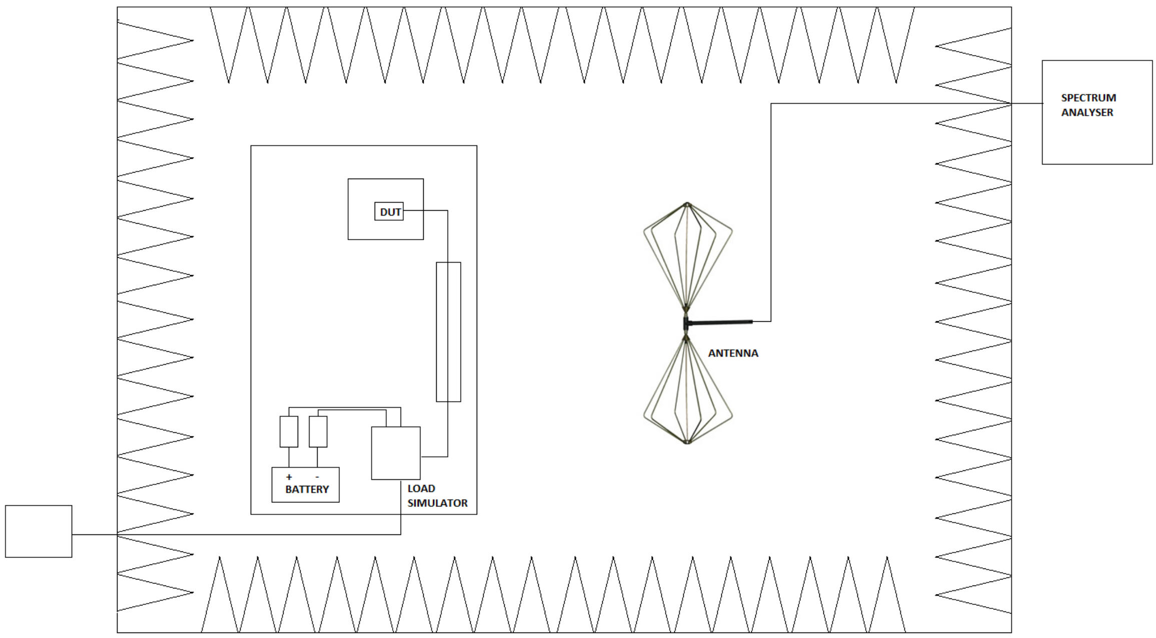


Figure 4.9. Radiated emission set up with Biconical antenna in HP configuration

As reported in the ISO11452-2 [8] the set up model is the one reported in the figure. Starting from the left there is a table that support all the required stuff. A battery is connected to two line impedance stabilization network LISN, both connected to ground, used as low pass filter between the power supply and the dut and to isolate it from the dut noise. The two lines than are connected to a load simulator that is normally used to simulate those situations in which the dut drive a load. They pass then on low relative permittivity $\epsilon_r < 1.4$ support and finally to the dut that is grounded locally. Finally the antenna is placed in the middle at a distance higher than 70 cm from the lines. It is connected outside to a spectrum analyser.

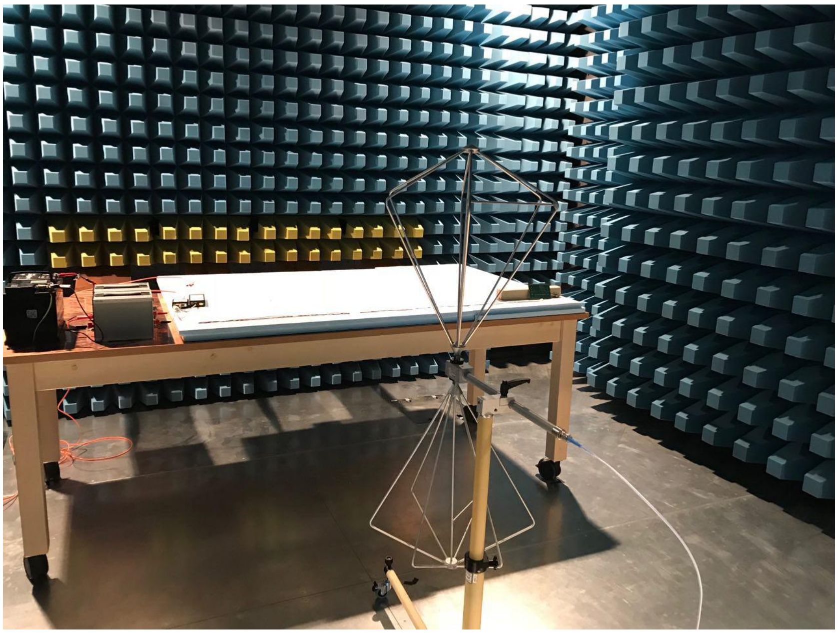
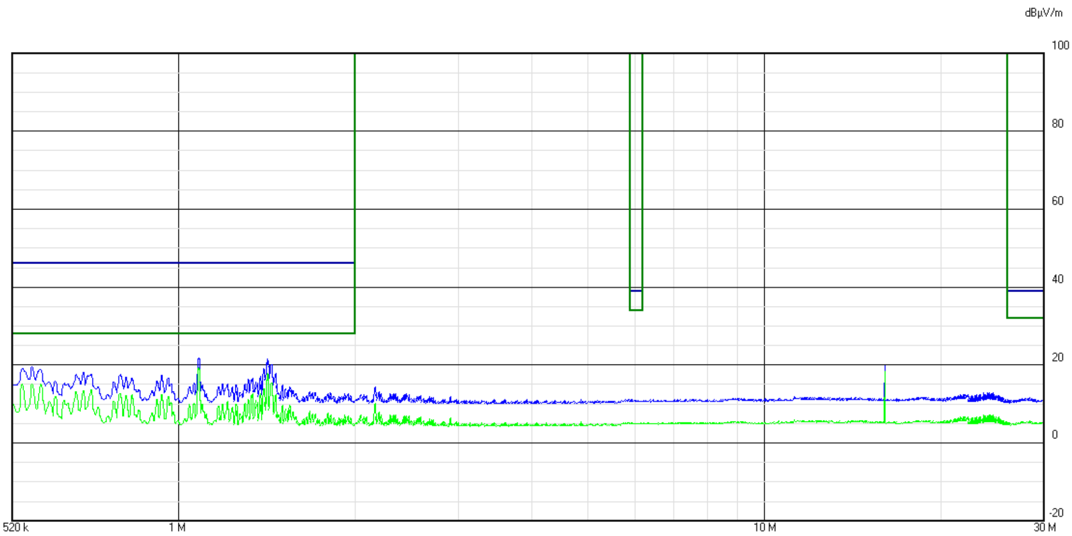


Figure 4.10. Radiated emission set up with Biconical antenna in VP configuration

The frequency range scanned with this test is in the range [0.52 : 2690] MHz and it is divided in multiple bandwidth that are analysed with different antennas and configurations.

Frequency range	**antenna**	**detection**	**bandwidth**	**step**	**polarization**	**position**
0.52-30 MHz	rod	avg, qPeak	9 kHz	2 kHz	HP,VP	central
30-300 MHz	biconical	avg, Peak	120 kHz	30 kHz	HP,VP	central
310-320 MHz	log-periodic	avg, Peak	9 kHz	2 kHz	HP,VP	central
430-438 MHz	log-periodic	avg, Peak	9 kHz	2 kHz	HP,VP	central
300-1000 MHz	log-periodic	avg, Peak	120 kHz	30 kHz	HP,VP	central
1000-2690 MHz	horn	avg, Peak	120 kHz	30 kHz	HP,VP	central

These rules are imposed by the FCA directive CS00054, that is the directive for all the electrical and electronic devices of a road vehicle, that defines the frequency ranges and also the limit at which the signal captured by the antenna must be submitted. The test is considered passed if all the conditions are satisfied. To be sure that the chamber is truly isolated by electromagnetic fields at the beginning of each test a measure of the background noise is executed, of course with the dut turned off. After this preliminary check the test has been performed, providing a pass result for all the requirements.



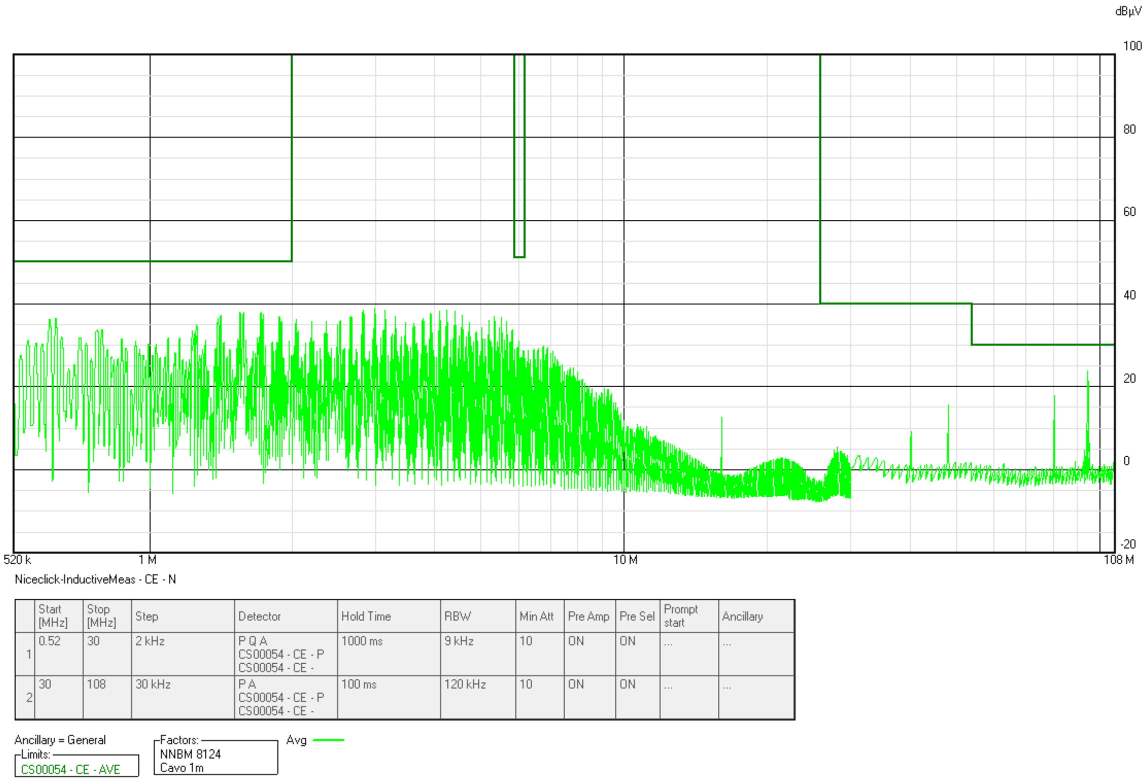
	Start [MHz]	Stop [MHz]	Step	Detector	Hold Time	RBW	Min Att	Pre Amp	Pre Sel	Prompt start	Ancillary
1	0.52	30	2 kHz	P Q A CS00054 - RE - P CS00054 - RE - A	1000 ms	9 kHz	10	ON	ON	...	...

Figure 4.11. Radiated emission result

Figure 4.11 show the result of measurement in the range 0.52-30 MHz and the limits defined by the directive. The green signal is relative to the average signal and as it is possible to see it is below the allowed limit, the same for the quasi-peak signal.

4.3.2 Conducted Emissions

The second test campaign concern the conducted emissions and relates to the mechanism in which an electronic device can generate electromagnetic energy and couple it to the power supply lines or communication ports. This means that if this happens the electromagnetic disturb is coupled to the entire power supply network causing problems to other connected devices. The range of frequency of this disturbs is normally lower than those investigated with the radiated emissions. The setup is the same of the previous test except for the antenna that in this case is not necessary , the screened room instead of the semi-anechoic and for the LISN configuration. In fact the measurement is done on one of the two artificial network with the other closed to an adapted load $Z_s = 50\ \Omega$, as described in the voltage method. The informations acquired on the LISN are then sent outside the chamber with a fibre optic transceiver.



	Start [MHz]	Stop [MHz]	Step	Detector	Hold Time	RBW	Min Att	Pre Amp	Pre Sel	Prompt start	Ancillary
1	0.52	30	2 kHz	P Q A CS00054 - CE - P CS00054 - CE -	1000 ms	9 kHz	10	ON	ON	...	...
2	30	108	30 kHz	P A CS00054 - CE - P CS00054 - CE -	100 ms	120 kHz	10	ON	ON	...	...

Figure 4.12. Conductive emission average signal result in 520 kHz-108 MHz range on N terminal

Figure 4.12 show the measurement results and the related limit that is respected. It is possible to see the variation of frequency step n the range 30-108 MHz as claimed in the directive. As before also this test has been passed.

4.3.3 Radiated Immunity

Next step is the first of two immunity test that aim to verify if the circuit is hardy enough when disturbs are injected on the power supply and communications line or irradiated. Radiated immunity check if the behaviour of the circuit is the same when an electromagnetic disturb is radiated on the dut. The test is performed in the semi-anechoic chamber in order that all the electromagnetic radiations are absorbed. The setup is similar to the radiated emission one but in this case the antenna is connected to a signal generator and then to a RF power amplifier able to provide the required power for the RF signal. The correct behaviour of the circuit can be seen in different ways, in this case the informations about the measurement are sent through a CAN bus toward an optic fiber transceiver. The transceiver bring the informations outside the chamber and then to the pc through a serial interface. This allow to verify in real time if the measure remain unchanged, since no force is applied to the sensor.

Frequency range	antenna	modulation	power	polarization	position
400-1000 MHz	quadrilog periodic	CW	80 Vm	HP	central
400-806 MHz	quadrilog periodic	AM	80 Vm	HP	central
806-1000 MHz	quadrilog periodic	PULSE A	80 Vm	HP	central
200-1000 MHz	quadrilog periodic	CW	80 Vm	VP	central
200-806 MHz	quadrilog periodic	AM	80 Vm	VP	central
1000-3200 MHz	horn	CW	80 Vm	HP	central
1200-1400 MHz	horn	PULSE R	100 Vm	HP	central
2700-3200 MHz	horn	PULSE R	100 Vm	HP	central
1000-3200 MHz	horn	CW	80 Vm	VP	central
1200-1400 MHz	horn	PULSE R	100 Vm	VP	central
2700-3200 MHz	horn	PULSE R	100 Vm	VP	central

These are the frequency range and modulations investigated. CW is un unmodulated sine wave that in the range 800-1800 MHz simulates the GSM disturbs. AM is a sine wave amplitude modulation that simulates radio interferences, while PULSE R and PULSE A are used to simulate radar/Wifi signals behaviour. Also in this case any problem is arisen.

4.3.4 Conducted Immunity: BCI

The final test is about conducted immunity and so the strength of the circuit when RF electromagnetic disturbances are present on the power supply lines and on the CAN bus lines. This is probably the most tricky test to pass. Bulk current injection uses a current probe, that is a coil, to induces RF disturbances to the power and communication lines in the form of parasitic currents. These parasitic currents induced in the CAN lines can for example alter informations and cause unwanted actions such undesired activations. The setup is the same of the conducted emissions but in this case a current probe is used. Of course the lines form the load simulator to the dut passes inside it but at two distances respect to the dut, 75 cm and 15 cm. The frequency ranges are lower than the radiated one due to the cut-off frequency of the coil. In fact the two ranges are 1-30 MHz and 30-400 MHz with AM and CW modulation. This test provide some problems to the circuit, in fact the raw value measured has some variations , however thanks to the baseline introduction the output force does not overcome the force present threshold and this allow to declare the test passed. This is due to the fact the disturbances on the power lines causes a V_{dd} variation that affect the charging phase allowing more or less power to be released in the circuit.

Chapter 5

Application: Force Touch Sensing

The initial idea of this project was to design and realise a force sensor system able to replace inaccurate sensor to monitor the force applied by the user over touch surfaces where haptic feedback is present. However at the end of the work I come up with the idea to use this system not only to monitor the force but also the touch sensing. Thus I started studying the forces applied to a surface[9] and their relative equations and I develop a Force Touch Technology able to monitor x, y, z coordinates over a plane surface with the only use of four haptic actuator placed at the four edges.

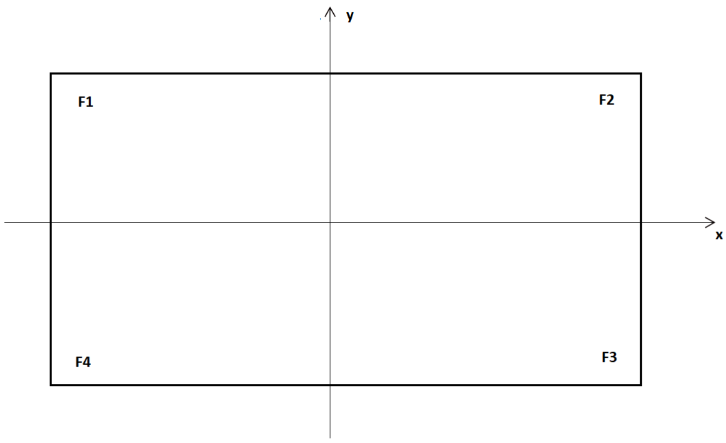


Figure 5.1. Touch force surface

The four actuators that provides also the force measurement as previously described are mounted on the four edges of the surface. Using force data and their relationship at each instant of time it is possible to compute the spatial coordinates as follow:

$$Position\ z\ = f1 + f2 + f3 + f4$$

$$Position\ x\ = \frac{100\ (f2 + f3 - f1 - f4)}{z}$$

$$Position\ y\ = \frac{100\ (f1 + f2 - f3 - f4)}{z}$$

These simple equations allow to determine a reliable estimate of the coordinates of the finger pushing the surface. In fact it compute the difference of the forces with respect to the center of the surface that correspond to the axis origins. Since it is possible that different measuring stages provides different resolutions and dynamics, due to the inductor and capacitor large uncertainties, the first step is to normalize all the forces measurement in a common range to correctly compare them.

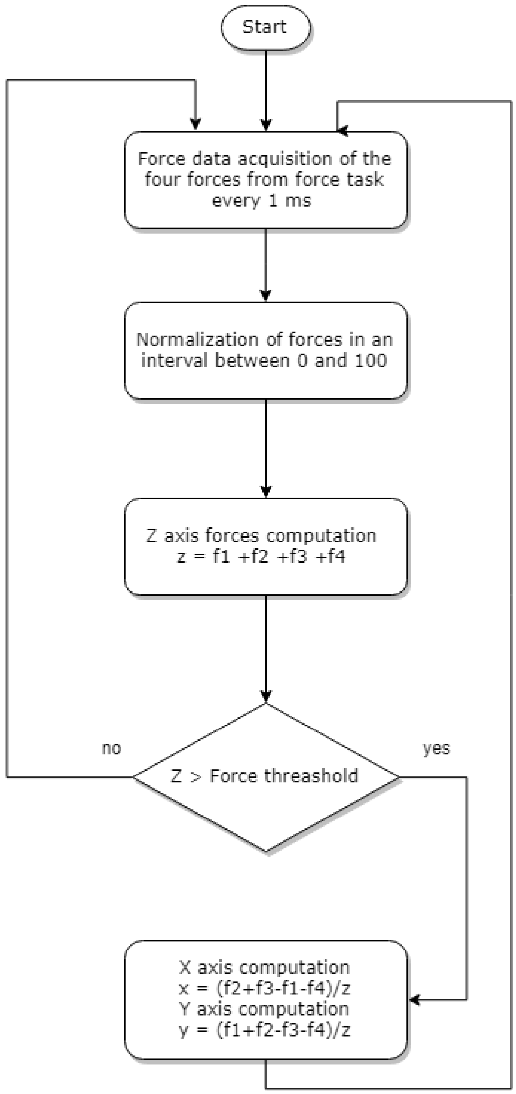


Figure 5.2. Touch coordinates algorithm

Then the first computation is about the total force applied and so the z component that is computed as the sum of all the forces. Higher the pressure higher the precision of the system, for this reason a force threshold is used. This solution allow to avoid noisy and inaccurate measurement. If the force applied is sufficient the X and Y coordinates can be computed as reported in the previous equations. These values are then normalized with respect to the total forces in order that the coordinates are not related with the force applied. To reduce high frequency noise present on the signals a low pass software filter has been used. In order to be able to use this solution for touch application an algorithm in charge of the recognition of user gestures on the surface is necessary.

5.0.1 Gesture recognition

Gesture recognition is a particular algorithm able to identify a particular gesture the user perform on the touch. There are five principal gestures that every system normally handle: the first two are finger movements though the y axis upward and downward that are called respectively *Scroll up* and *Scroll down*. The motion through the x axis instead are called *Swipe left* and *Swipe right*. Finally the *Tap* is a gesture that detect a finger touching the screen for a short time, almost 100 ms, and it is used to select/activate an object or a function in a common graphic interface.

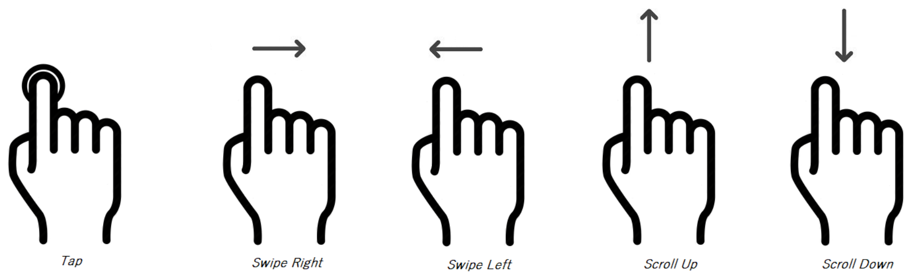


Figure 5.3. Gesture types supported

To handle with this function another task is created in order that every 5 ms it receives data about the spatial coordinates and checking the track crossed by the finger can identify the correct gesture. The algorithm starts when the z axis coordinates overtake a force threshold. Here the initial coordinates x and y are stored and they will be used later to compute the maximum displacement. Then at each iteration the algorithm check for the maximum value in module reached in the plane and finally when the finger is released and the force condition is no more valid the gesture can be computed.
A first check about the time duration of the gesture is needed to sort the Tap from the other. If the time span is higher than 100 ms then Δx_{max} and Δy_{max} are computed and according to their relation it is possible to sort. An higher through x axis means a Swipe gesture is detected and to select the correct direction, left or right, the sign is needed.

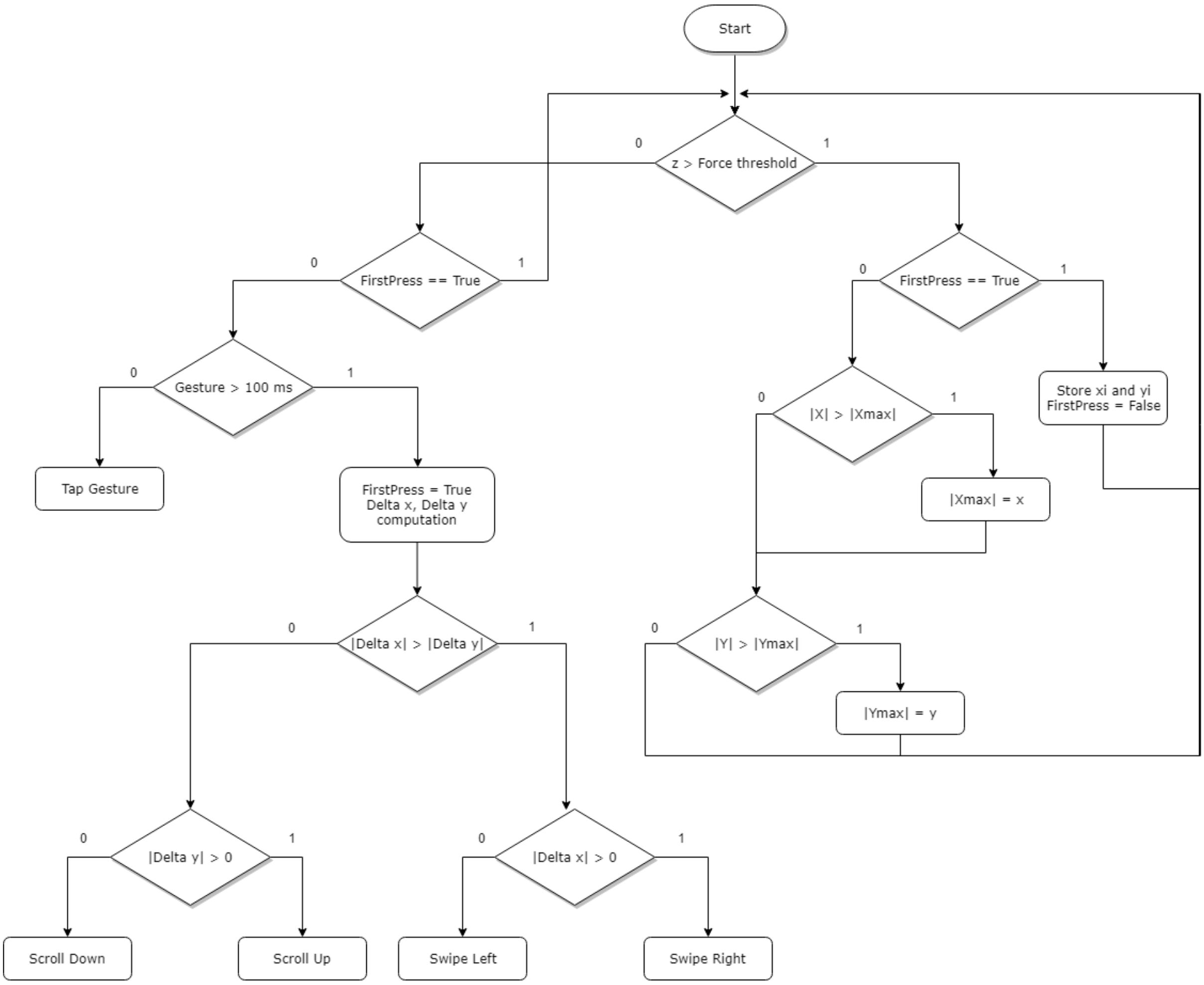


Figure 5.4. Touch Recognition algorithm flowchart

The algorithm flowchart is reported above. This solution includes the most common gestures required from a touch device to let the user navigate in a context menu for example. As previously said the coordinates algorithm works well if the forces applied are higher than a certain threshold that has been measured using a dynamometer and it is about 2 Newton. This threshold is higher than the force normally the user is used to apply to a touch surface, that is about 0.1-0.5 Newton. However this condition is related to the actuators elastic component, the O-ring, which has an high elastic constant. So to reduce the this threshold it is sufficient to drop the elastic constant by changing the material for example or by choosing an O-ring with a lower hardness.
Fig 5.5 report the acquisition of coordinates data acquired while testing the gesture algorithm. The first signal to rise is the z coordinate, the force, that let the system wait for the exceed of the threshold to start computing also the X and Y coordinates. After 200 ms, time required to reach the threshold, the X start rising while the Y stay stable around the zero. This is interpreted by the algorithm as a *Swipe right*.

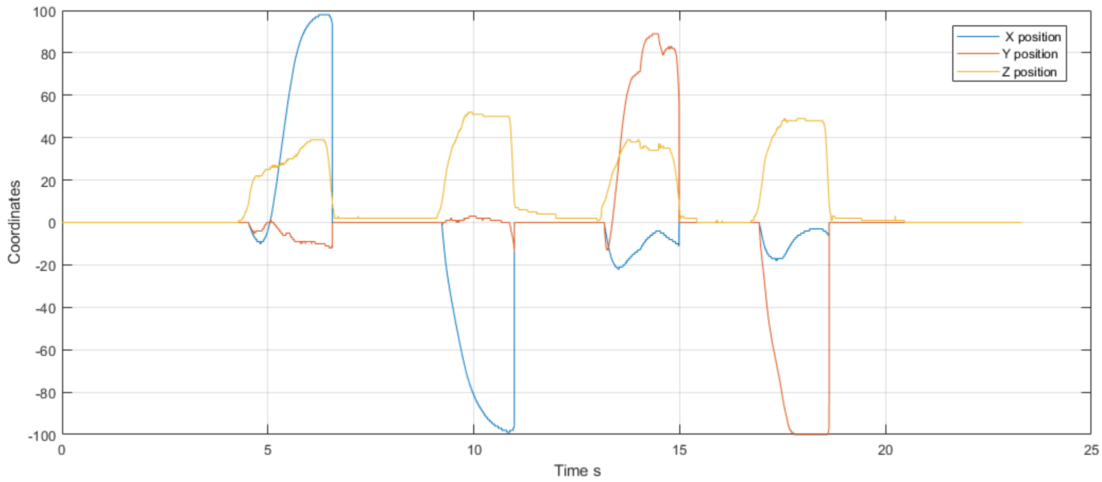


Figure 5.5. Spatial coordinates computed by Fource Touch Algorithm

To better study the goodness of the algorithm a series of acquisitions has been performed and a statistic of the errors has been realized. Two different threshold has been studied, 2 and 5 Newton, to compare them and understand which solution is better overall.
Before starting analysing data collected a consideration must be taken into account regarding the habit of the user. In fact humans and their senses are methodical and since our fingers are used to apply a certain quantity of force on a touch screen, at their first use of this solution the error is higher.

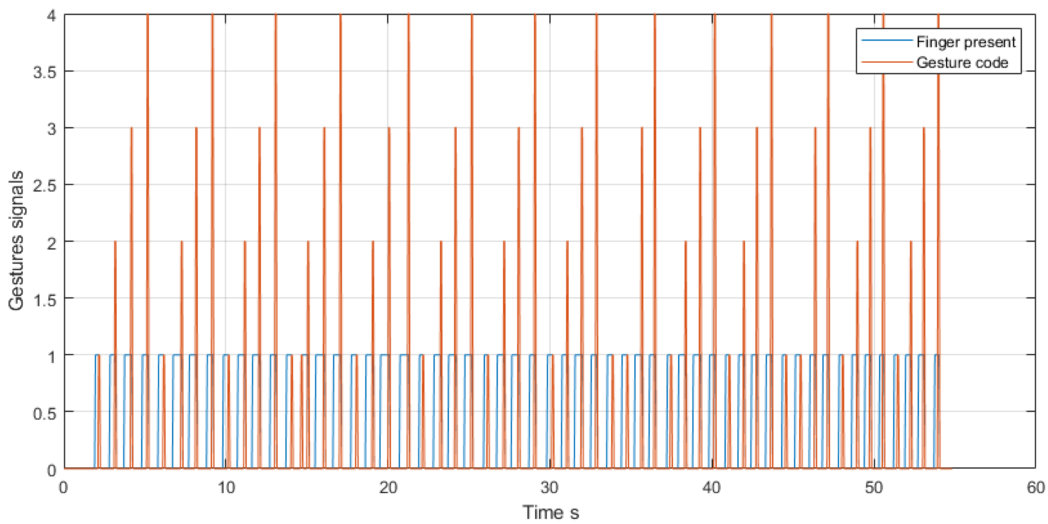


Figure 5.6. Gesture Recognition Algorithm error rate using 2 Newton force threshold

However with a small practice the user accustom its hands and the error drops down. For this reason two tests are performed: a first one without any type of practice of the user and the second one after some minutes of tries. These two test provides different results for the reasons explained before. The first use test provide a statistical error of 15-20 % while the practise test provide an error about 5% with the lower threshold and a 2 % error with the higher one. Another observed phenomena is that the introduction of the haptic feedback and the relative sound the error of the user decrease further because the user knows if the force is sufficient or not through the feedback.

Chapter 6

Conclusion

This project started because of the need of a new force sensor able to provide a precise and repeatable output that should be used to improve the reliability of a touch and haptic interface. The previously adopted sensors however were inaccurate, expensive and with mechanical limits. From this need a new sensorless system has been conceived, realized and finally tested.
This allowed to replace an old and imprecise force sensing system with this one opening new technological solutions. In fact due to its high precision and its high dynamic range it is possible to develop new touch system where is no more required the use of noisy capacitive touch but only to read forces from different points, the same from which the haptic feedback come from, and from these data providing the localization also on x-y axes. This final step allow each surface to became touch and haptic with the use of more than two actuators and an electronic board, strongly reducing the cost of production.

Bibliography

[1] N. Shin, "*Touch Sensors vs. Mechanical Buttons*", www.silabs.com/community/blog, 2014

[2] https://www.dmccoltd.com/english/museum/touchscreens/technologies/Projected.asp

[3] Culbertson H.,Schorr B.S., Okamura M. A.,"*Haptics: The Present and Future of Artificial Touch Sensation*", Annual Review of Control, Robotics, and Autonomous Systems,Vol. 1, pp.385-409, 2018

[4] El Saddik A. , "*Haptics Technologies*", Springer Series on Touch and Haptic Systems, 2011

[5] Zwislocki J.J., "*Sensory Neuroscience: Four Laws of Psychophysics* ", Springer, 2008

[6] McClelland C. "*Intro To Haptic Technology*", www.iotforall.com, 2013

[7] Gogue P. G. , Stupak J. J. "*Theory & Practice of Electromagnetic Design of DC Motors & Actuators* " , G2 Consulting, Beaverton

[8] ISO 11452-2, "*Road vehicles — Component test methods for electrical disturbances from narrowband radiated electromagnetic energy —*", second edition 2014

[9] Frish J.K., Faggin F., Leyh G.E., Platt J.C., Allen T.P., Schediwy R.R.,"*Force Sensing Touchpad*", Patent Number: 5,854,625,1998