Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI ,13), October 28–30, 2013, Eindhoven, The Netherlands.

Opportunistic Synergy: a Classifier Fusion Engine for Micro-Gesture Recognition

Leonardo Angelini[1,2], Francesco Carrino[1,2], Stefano Carrino[1,2], Maurizio Caon[1], Denis Lalanne[2],Omar Abou Khaled[1], Elena Mugellini[1]

[1]University of Applied Sciences Western Switzerland, Fribourg

[2]University of Fribourg, Switzerland

[1]{firstname.lastname}@hefr.ch, [2]{firstname.lastname }@unifr.ch

ABSTRACT

In this paper, we present a novel opportunistic paradigm for in-vehicle gesture recognition. This paradigm allows using two or more subsystems in a synergistic manner: they can work in parallel but the lack of some of them does not compromise the functioning of the whole system. In order to segment and recognize micro-gestures performed by the user on the steering wheel, we combine a wearable approach based on the electromyography of the user's forearm muscles, with an environmental approach based on pressure sensors integrated directly on the steering wheel. We present and analyze several fusion methods and gesture segmentation strategies. A prototype has been developed and evaluated with data from nine subjects. The results prove that the proposed opportunistic system performs equal or better than each stand-alone subsystem while increasing the interaction possibilities.

Categories and Subject Descriptors

H.5.2 [**Information Systems**]: Information Interfaces and Presentation; I.2.M **[Computing Methodologies]**: Artificial Intelligence – Miscellaneous;

Keywords

Environmental and Wearable paradigms; Ubiquitous computing; Tangible Gestures; Electromyography; Micro-gestures; In-Vehicle Interaction.

1. INTRODUCTION

In the last decades, humans established a profound relationship with technology arriving to a symbiotic relation. This trend, called *human-computer confluence*, aims at enhancing the user experience introducing new forms of sensing, perception, interaction, and comprehension [1]. In the ubiquitous computing era, technology is pervasive and is embedded in the environment and in all the objects of everyday life. In particular, the cockpit of the vehicle is undergoing important technological development. In this case, technology aims at enhancing both safety and user experience. However, the increase of interactive systems has the serious drawback of overwhelming the user with excessive information to handle [1]. This can lead to the distraction from the primary task, i.e., driving. Natural interaction addresses this issue enhancing the user experience while reducing the cognitive load induced by secondary tasks. In particular, gestural interactions represent "promising means to cover the full range of a driver's operational needs while minimizing the cognitive and visual workload", as stated by Riener in [2].

AutomotiveUI '13, October 28 - 30 2013, Eindhoven, Netherlands Copyright 2013 ACM 978-1-4503-2478-6/13/10…$15.00. http://dx.doi.org/10.1145/2516540.2516563

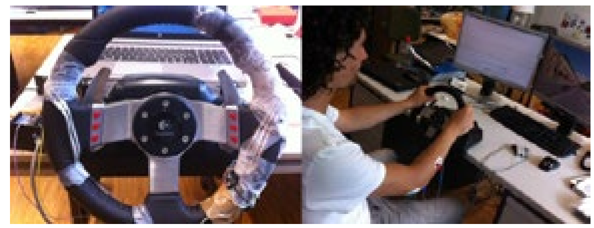

Figure 1. Left: environmental sensors. Right: wearable sensors.

Ubiquitous computing encompasses two paradigms: environmental and wearable computing. The environmental computing paradigm uses distributed sensors in the interaction space to detect and react to the inhabitant's movements, gestures, and activities. The wearable computing paradigm uses sensors worn by the person for detection and sensing. Environmental and wearable computing paradigms have different and complementary advantages and drawbacks [3]. In this paper, we present a hybrid approach that opportunistically combines the environmental and wearable computing in a synergistic manner. In fact, the aim of our system is to combine the strong features of both environmental and wearable computing while alleviating their weaknesses.

Therefore, the main contribution of this work is the definition and the investigation of an opportunistic paradigm that supports gesture-based natural interaction in the car. In particular, we analyze three aspects:

- Design parameters for the interaction (advantages and limitations).
- Classifier fusion engine.
- Manual and automatic gesture segmentation.

Gesture is the interaction modality taken into account in the system. In particular, we focus on micro-gestures [4] that are economic, subtle and require reduced effort from the user. The natural interaction studied in this paper focuses on a specific scenario: the use of an In-Vehicle Information and communication Systems (IVIS) while driving. Users perform micro-gestures on the steering wheel to interact with the IVIS.

Five gestures control the car stereo switching it on and off, switching channels and controlling the volume. We investigate a recognition approach based on pressure sensors disposed on the steering wheel (Figure 1 left) as environmental component;

wearable sensing is based on electromyography (EMG), with electrodes positioned on the driver's arms (Figure 1 right).

Section 2 situates our work in the state of the art. Section 3 introduces the proposed paradigm and the investigated methods for classifier fusion and gesture segmentation. Section 4 describes the system implementation. Section 5 evaluates and discusses the performances of the system. Finally, Section 6 concludes the paper.

2. STATE OF THE ART

An emerging trend in automotive research is designing and developing natural human-computer interaction systems [5]. In particular, gestural interfaces captured the attention of the research community since they can improve the user experience and increase safety [2]. Indeed, Gonzalez et al. stressed the importance of designing gestural interfaces that help the user to keep "eyes on the road and hands on the wheel" [6]. Their proposal consisted in a system that enabled the user to input text through small thumb gestures while holding the steering wheel. Following a similar concept, Döring et al. integrated a multi-touch interactive screen in the steering wheel in order to allow the user to interact with both thumbs [7]. A different approach but with the same goal has been presented by Mahr et al. in [8]. Their system, called Geremin, recognizes 2D micro-gestures performed with the index finger while the rest of the hand holds the steering wheel. The WheelSense system, presented in [9], shows another interesting concept: the micro-gestures are designed to keep the whole hand on a safe position of the steering wheel even while interacting with the IVIS.

The aforementioned examples of gestural interfaces involve only systems with sensors integrated in the environment, i.e., the steering wheel. Other systems that do not adopt the environmental paradigm exist. For instance, the work presented in [10] explores the possibility of an electromyography-based system that recognizes micro-gestures performed with the fingers. In this case, the sensors are distributed on the user's arm. In fact, this work moves towards a wearable paradigm even if it is not a pure one. A similar project, described in [11], proposes also an architecture based on the connection between the wearable part of the system and the IVIS. Such architecture distributes the gesture recognition process between the wearable and the environmental computing systems. Other works adopted this hybrid approach coupling wearable systems (smartphones), with environmental systems (the IVIS). An example is the Terminal Mode project, where the smartphones are integrated in the IVIS attaching them to a dock in order to enable users to always have their personal applications and services at their fingertips [12]. In fact, this concept takes advantage of one of the most important features of wearable computing: personalization.

The importance of mixing wearable and environmental computing has already been highlighted by Rhodes et al. in [3]; they presented a peer-to-peer network of wearable and environmental systems, which mixes the complementary advantages coming from both paradigms. To be more precise, this architecture aims at bringing together privacy and personalization provided by the wearable paradigm with localized information, localized control and resource management characterizing the environmental paradigm. The work presented by Carrino et al. in [13] wanted to extend the previous concept of mixing together the benefits driven from the adoption of these paradigms introducing also the consistency feature. The architecture proposed by the authors focuses on gesture recognition and deals with the features and classifier results fusion.

Extending the concepts developed in these last two works, we propose an opportunistic paradigm in order to exploit complementary advantages, to increase accuracy and to enhance interaction experience.

3. OPPORTUNISTIC SYNERGY

This section describes and analyzes the proposed opportunistic paradigm and its underlying challenges: the multiclass classifier fusion and the manual and automatic gesture recognition and segmentation.

Definition: *an opportunistic multi-component system is synergistic if all the components contribute to the overall system performances, but the lack of some of them does not compromise the proper functioning of the whole system. A system with multiple components should perform better than its individual components. Performances are measured in terms of interaction, usability or accuracy.*

In our specific case, we have two components: the environmental (the pressure sensors on the steering wheel) and the wearable (the EMG sensors on the user's arms). Our system combines the complementary advantages of these two components. Users not wearing sensors can interact with the environment according to the "*come as you are*" design [14]. On the other hand, users wearing sensors can exploit them to interact with the system without the need of an augmented steering wheel. Finally, if sensors are present both in the environment and on the user a richer and more accurate interaction can be performed.

3.1 Design Parameters

The environmental and the wearable computing can be designed through eight parameters:

Interaction area is the space in which the user interactions and commands are sensed by the system.

Personalization is the capacity of the system to provide and maintain personalized interactions for the users.

Consistency is the capacity to improve the system thanks to the prolonged, continuous interaction between the human and a computer.

Private interaction and intimate interfaces are "*discre[e]te interfaces that allow control of mobile devices through subtlety gestures in order to gain social acceptance*" [15].

Localized information is the system feature that specifies how to access the information in a specific location (such as the cockpit, the windshield, or the steering wheel).

Localized control is the system feature that specifies how and where to provide information and commands to the system.

Resource availability is strictly linked to the current technologies adopted for the interaction, e.g., processing power, energy, etc.

Resource management is the system capability to efficiently handle the different available resources. A smart environment, like a vehicle of the next future should deal with heterogeneous sensors, actuators, and multiple users with different tasks and needs.

As Rhodes et al. claimed [3], the wearable paradigm is advantageous to facilitate private interaction and the management of personal information; for example, it is simple to define a personalized profile for each user.

The environmental paradigm strong points are the localized information, the localized control and the resource management. Localizing the control and the information in the environment can impose physical constraints. These constraints can be facilitators to the interaction or can improve safety. For example, in our prototype (detailed in Section 4.1), we placed the 5 pressure sensors in specific regions of the external ring to force the user to be compliant with the hands position suggested by the Swiss driving school manual [16].

The system discussed in this paper is depicted in Figure 2. The dotted connectors represent loose links.

In order to realize a synergistic paradigm, the fusion of the wearable and pervasive blocks should be accurately designed.

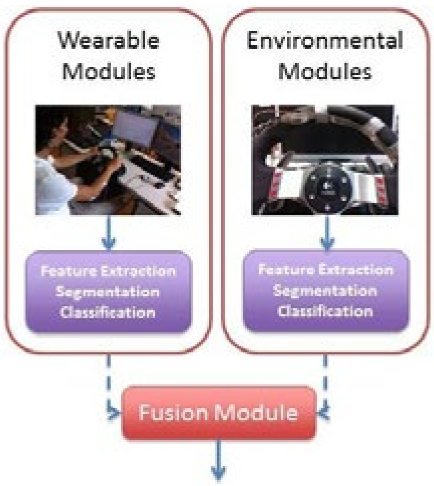

Figure 2. Synergistic paradigm in the car. Sensors are embedded in the steering wheel and worn by the user.

3.2 Classifier Fusion Engine: Methods

The fusion of multiclass classifiers is a critical task in machine learning problems. In particular, the sub-problem of performance measures of multi-class classifiers is still an open research topic [17].

The fusion of the data coming from the wearable and pervasive classifiers is crucial in order to profitably merge the information and improve the performance of the synergistic system. Our schema is based on a *late fusion approach* and we investigate different methodologies to score and weight the results of each class probability.

The late fusion approach has the advantage to be less affected by changes and modifications in a subsystem and, therefore, facilitate the realization of a *synergistic paradigm.*

The fused decision is selected through the Sum rule. As Kittler et al. [18] defined: "*The Sum rule operates directly on the soft outputs of individual experts for each class hypothesis, normally delivered in terms of a posteriori class probabilities. The fused decision is obtained by applying the maximum value selector to the class dependent averages of these outputs.*"

Kittler et al. [19] also stated that, generally, "*the combination rule developed under the most restrictive assumptions – the sum rule – outperforms other classifier combination schemes.*"

In a two-classifier problem the sum rule is reduced to the comparison of the output of the two experts. It is important to note, that the term *"expert"* encompasses both the classifier and the subsequent weighting process.

In this research, we conceived a classifier fusion engine in order to compare 10 different approaches estimating the soft output of the experts. The weighting process is based on confusion matrices and the classifiers soft outputs that we have estimated in a cross-validation phase (performing a k-fold cross-validation on the training set).

The following subsections present the weighting processes that we have studied; some of them come from the literatures (the Sum Rule and the naive Bayes combination) and are used as comparison for our propositions.

3.2.1 Sum Rule (SR)

This method uses directly the likelihood outputted by the classifiers as weights for the Sum rule [19].

3.2.2 Naive Bayes combination (NB)

The NB combination [20] is very common in literature as decision rule for the fusion of classifiers because it could be easily extended to more than two classifiers. It exploits the conditional a posteriori probability $P_j(i|s_j)$ of each classifier that the received gesture belongs to the i^{th} class, given that the j^{th} classifier has assigned that gesture to the s^{th} class (s_j). These conditional probabilities can be calculated by the confusion matrix *CM* generated for each classifier during the training phase. Under the assumption that the classifiers are mutually independent, the NB combination calculates the overall a posteriori probability of a class *i* as:

$$W_{Bayesian} = P(i) = \prod_{j=1}^{L} \frac{cm_j(i,s_j)}{\sum_{k=0}^{C} cm_j(k,s_j)}, \qquad i=1,..,C \qquad (1)$$

Where L is the number of classifiers, $cm_j(\cdot,\cdot)$ are the confusion matrix elements of the j^{th} classifier, and C is the number of classes.

From the NB definition, we can see that if a classifier emits a soft output this value is not taken into account by the NB classifier. For this reason, we include in our tests also a modified variant of the NB approach, consisting in the multiplication of *P(i)* by the soft output of the classifier (e.g., the likelihood in the case of HMMs). In the rest of the paper, we refer to this method with the abbreviation *NBW*.

3.2.3 Matthews Correlation Coefficient method

The Matthews Correlation Coefficient (MCC) [21], also known as *φ-coefficient* in the binary case, is an aggregate objective function (AOF) that is used in machine learning as performance measure in the context of multiclass classification. Since this coefficient is stable even if classes are of different sizes, it is adapted to our multi-class problem using a one-vs-all approach (see [22] for a deeper insight on the use of MCC in machine learning). MCC is defined as:

$$MCC = \frac{(TP * TN - FP * FN)}{\sqrt{(TP + FP) * (TP + TN) * (FP + FN) * (TN + FN)}} \qquad (2)$$

Where TP indicates true positive, TN true negative, FP false positive, and FN false negative. This correlation coefficient returns a value between -1 and +1. A value of -1 indicates total disagreement between the classifiers, 0 is for completely random prediction and +1 means total agreement.

The MCC of the aggregate confusion matrix (aggregation of k matrices obtained through the k-fold approach on the training set) is used as weight in the MCC method. The MCC is computed for each gesture class.

To solve potential disagreement between the two classifiers, it is important to introduce in the fusion process the information about the confidence of a classifier in predicting the class guessed by the other classifier. For example, if the first classifier predicts the class s_1 and the second the class s_2, it is important to know the performance of the first classifier on the class s_2 and of the second classifier on the class s_1. In order to deal with this information, we conceived the *MCC Conditional* approach (abbreviated as MCC+). Finally, in a configuration with 2 classifiers, the weight of a classification result is computed as the multiplication of the MCC of the predicted class and the MCC of the class predicted by the other classifier:

$$W_i^{j_1} = MCC_{s_1}^{j_1} * MCC_{s_2}^{j_1} \quad (3)$$

Where the notation $MCC_{s_j}^{j}$ indicates the MCC coefficient computed on the confusion matrix of the classifier j predicting the class s_j. In particular, in equation 3: s_1 is the class predicted by the classifier j_1 and $s_2$is the class predicted by the classifier j_2.

The main drawback of this approach is that it needs to be reformulated in order to work with more than two classifiers.

3.2.4 Schemes variant

For all the previous approaches we have designed a variant approach taking into account the overall accuracy of the classifiers after a cross-validation step. Multiplying a weight by the classifier overall accuracy brings the fusion algorithm to increase the confidence on the classifier that perform better during the cross-validation phase.

For example, $W_{Bayesian*}$ is computed by multiplying the base approach (i.e., BN) by the overall accuracy. Then:

$$W_{Bayesian*} = W_{Bayesian} * Overall\_Accuracy \quad (4)$$

3.3 Manual and Automatic Gesture Recognition and Segmentation

For the realization of a synergistic paradigm a critical step is the processing of the segmentation signals. Mainly for two reasons: firstly, from an interaction point of view the segmentation can have an impact on the cognitive load on the user; secondly, having sub-systems working asynchronously, the fusion system can receive gestures that are delayed and it can deal with missing signals. We addressed the possible delay between the outputs of the different classifiers by fixing a maximum delay of 500ms to be considered as belonging to the same gesture. Otherwise, gestures are treated separately and results are not merged. This approach implies a minimum delay (the 500ms) between the user gestures.

To deal with the cognitive load on the user, we compared two approaches. The first, called manual segmentation, requires the user to communicate to the system the start and the end of a command gesture. This approach helps the system to understand if a gesture is intended for interaction with system or it is a manipulation of the steering wheel due to the normal driving. A simple way to achieve the manual segmentation is to perform a segmentation gesture on the steering wheel with the left hand while the right hand is performing the actual command gesture. The second, called automatic segmentation, implies that the algorithm used to recognize the gesture can automatically detect if the gesture is intended for the IVIS or it is normal steering wheel manipulation for driving.

3.3.1 Manual Segmentation

Manual segmentation involves the use of the left hand in order to directly provide the system the *start recognition* and the *stop recognition* commands.

A practical example (used in our prototype) is the following: while the user is squeezing the steering wheel with the left hand, a gesture performed on the steering wheel by the right hand is meant as command and must be recognized.

Figure 3 presents the architecture of the system based on manual segmentation.

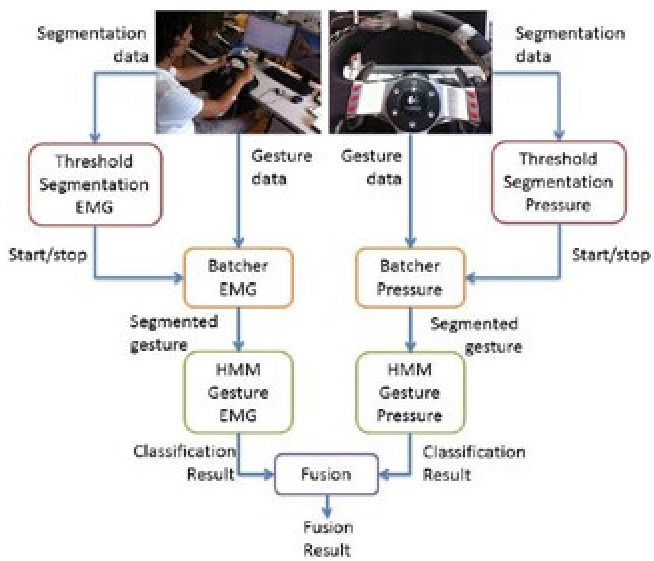

Figure 3. Manual segmentation architecture.

The sensors provide different types of information flow: segmentation data and gesture data. Information from EMG and pressure sensors is analyzed in parallel.

The first step involves the processing of the segmentation data. The data is segmented using thresholds that are fixed in a user-dependent calibration phase. In particular, we follow a hysteresis-based approach, in which the activation and deactivation thresholds are calculated taking into account the standard deviation of the user's signals when she/he is driving or she/he is performing the *Squeeze* (see Figure 4). Therefore, the segmentation is enabled when the signal (the root mean square for the EMG and the raw data for the pressure sensor) exceeds the *activation threshold* and is disabled when the signal drops below the *deactivation threshold*.

The Batcher modules start to accumulate the data on the start signal. With the stop signal the collected information is sent to the classification algorithm.

The Hidden Markov Model (HMM) classifiers were configured with 4 hidden states with forward topology and implementing the Baum-Welch algorithm to find the unknown parameters.

Finally, the outputs of the HMMs are merged in the fusion module.

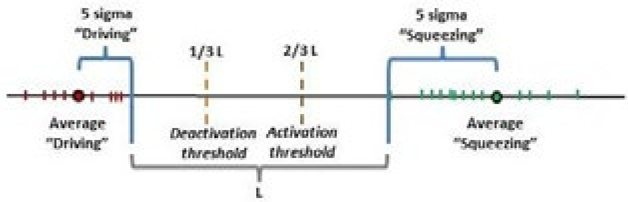

Figure 4. Manual segmentation - Threshold computation.

3.3.2 *Automatic Segmentation*

The automatic segmentation uses the same signals for gesture recognition and segmentation. Practically, it involves a reduced cognitive load on the user who has to perform only right hand gestures on the steering wheel.

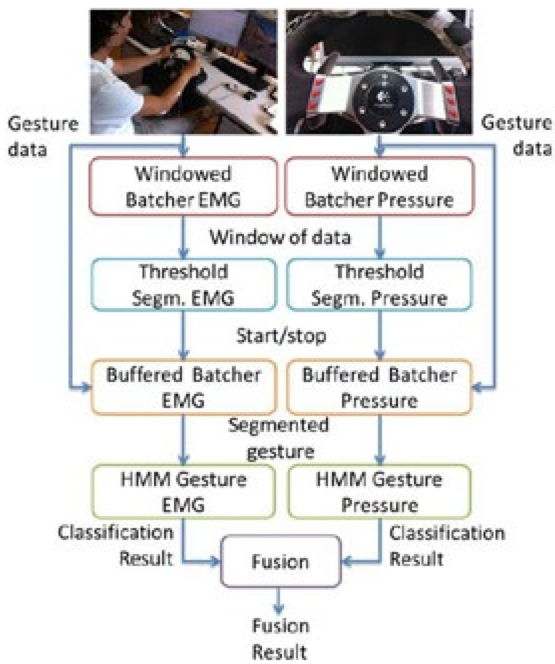

Figure 5. Automatic segmentation architecture.

The architecture, presented in Figure 5, is slightly different from the one presented in Figure 3. In this configuration only gesture data are used. In order to spot the start and the end of a gesture the system analyzes the data flows using a windowing approach.

For each window a SVN classifier estimates if the user is performing a gesture or is simply driving. The SVN algorithm uses a Gaussian kernel in which an appropriate value of sigma is obtained using the approach explained in [23] as implemented in the Accord.Net framework [24].

Since the segmentation events are fired after a window of data, there is the risk of losing important information. The buffered batcher modules allow queuing data and avoiding this loss.

The following part of the schema is the same as in the manual segmentation approach, with HMMs classifiers analyzing the data for gesture recognition.

3.3.3 *OR, AND and ADAPTIVE Segmentations*

We designed three different segmentation strategies in order to analyze the influence of the segmentation on the synergy. The OR strategy considers as valid results the contribution of both single and coupled classifiers (as the logic OR). The AND strategy takes into account only gestures simultaneously detected by the two classifiers. The ADAPTIVE strategy opportunistically switches between the OR and the AND strategies according the following rule using the confusion matrices of the segmentation:

If the classifier j_1 segments a gesture s ignored by the classifier j_2, and if $FP_{J_1}(s) > FN_{J_2}(s)$ then the detected gesture is considered a FP, i.e., incorrectly segmented. Otherwise, the gesture is considered as correctly segmented.

3.4 Micro-Gesture Vocabulary

The micro-gesture vocabulary has been designed to help the user to keep attention on the road and the hands on the steering wheel as suggested by Wolf et al. in [4]. To be more precise, this vocabulary comprises five tangible micro-gestures as depicted in Figure 6. In fact, these micro-gestures are designed to be performed while holding the external ring of the steering wheel exploiting its physicality in order to provide haptic feedback, which turns the wheel into a tangible interface as explained in [9].

Figure 6. The proposed micro-gesture vocabulary.

4. SYSTEM REALIZATION

4.1 Prototype

Figure 1 depicts the setup of the prototype that we realized in order to validate our proposition. As previously presented, our configuration is composed of environmental and wearable components. The environmental system exploits pressure sensors mounted on the external ring of a steering wheel; the wearable system is based on EMG electrodes positioned on the left and right user's arms.

4.1.1 *Environmental Components*

The pressure sensing system is based on five Tekscan FlexiForce sensors with a working range of 0-1lb [26] placed on a Logitech G27 Racing Wheel. Sensors are connected to an Arduino Duemila board. Data are digitalized and then sent to a PC at a frequency of 50Hz. For the positioning of the sensors on the steering wheel refer to Figure 1. Further details on this system can be found in our previous work [9].

4.1.2 Wearable Components

Sensors were positioned by non-medical personal following Cram's guide [27]. On the left arm, the sensor used for segmentation was placed on the *Flexor Carpi Ulnaris*. On the right arm we used the electrical activity of the *Extensor Carpi Ulnaris Flexor Carpi Ulnaris*, *Extensor Digitorum*, *Flexor Carpi Radialis*, and *Palmaris Longus* (Figure 7). The possibility to easily integrate the electrodes in clothes or in an armband dictated the choice of the muscles, which excluded the hand muscles. For each sensor, we extract the following features: signals Root Mean Square, Logarithmic Band Power and Mean Absolute Value, for a total of 12 features for the right arm and 3 features for the left. For the setup we used a Marq-Medical MQ16 device. Further details on this system can be found in our previous work [11].

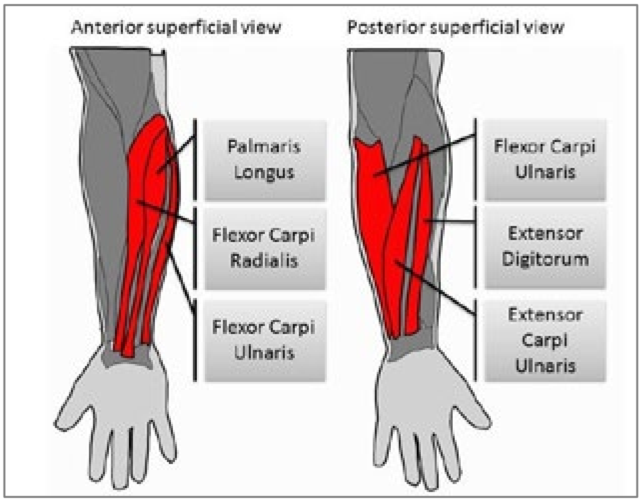

Figure 7. Highlights of the used muscles.

4.1.3 Parameters for Automatic Segmentation

In the configuration for automatic segmentation we could configure the window length and the overlapping percentage. To achieve the results presented later, we used a window size of 3 and 10 samples with an overlap of 33% and 50% for the EMG and the pressure data respectively. These values have been selected empirically during a pretest phase.

4.1.4 Software Framework

For the machine learning implementation we used a framework we developed in the past years [24], based on the *Accord.Net* framework [25].

4.2 Acquisition Protocol

We tested the system with 9 users (1 female), aged 24-31. After about five minutes for the familiarization with the system, we proceeded to the *calibration phase*. For this phase we adopted a user-dependent solution, which has the advantage to generally perform better than user-independent solutions. The drawback is that the system will always require a calibration and a training phase on the first usage. In order to mitigate these factors we designed a system that needs reduced training data. The calibration phase aim is to detect the average strength applied by the user on the steering wheel and the average electrical activity of the muscles. Such signals are user-dependent, depending on the user driving habits and muscular development (with significant differences between genders).

As next step, each user was asked to perform 30 gestures for each of the five gestures (*Squeeze*, *Up*, *Down*, *Tap*, and *Push*) in a random order. During the execution of the gestures, the user was asked also to perform normal driving actions on the steering wheel: turn left, right, accelerate, brake or stay in a rest position.

The acquisition protocol consists of the simultaneous acquisition of data for the manual and automatic segmentation. We asked the users to squeeze the external ring of the steering wheel with the left hand to perform the manual segmentation. This information was then used only by the manual segmentation modules. The whole acquisition process was performed in two sessions of about 15 minutes in order to allow the users to rest.

The 70% of the gestures were used as training and cross-validation sets; the 30% of the gestures were used as test set. We used a k-fold (k=10) cross-validation on the training in order to calculate the confusion matrix and to select the weights for the fusion module.

5. SYNERGISTIC PARADIGM: EVALUATION AND DISCUSSION

The goal of the prototype is to provide data to evaluate three aspects of the proposed synergistic paradigm: the interaction opportunities and limitations, the sensor fusion methodologies and the gesture segmentation.

Since the interaction opportunities are strictly related to the results of the fusion and segmentation steps, a quantitative evaluation of data fusion and gesture segmentation strategies is presented. Subsequently, a qualitative discussion about the interaction features of the synergistic paradigm is provided at the end of this section.

5.1 Evaluation of the Classifier Fusion Engine and Segmentation

We based our analysis on the comparison the F_1-score and the accuracy in the classification. The F_1 score is defined as:

$$F_1 = 2 * \frac{precision * recall}{precision + recall} \tag{5}$$

Where

$$precision = \frac{TP}{TP + FP} \tag{6}$$

$$recall = \frac{TP}{TP + FN} \tag{7}$$

The accuracy is defined as the number of TP divided the total number of gestures.

Table 1 presents these values for the different classifier fusion methods and segmentation strategies. Pressure and EMG rows present the results of the single classifiers. We can observe that in our configuration the EMG performs generally better than the pressure sensors.

The best results have been achieved with the ADAPTIVE segmentation strategy and the MCC* fusion method. Such configuration performs better of the other fusion methods and, in particular, equal or better than the best standalone classifier.

The proposed fusion approaches are independent from the classifiers typology and can be extended to classifiers with crisp outputs (the Sum rule is then reduced to a Vote rule). However, when the fusion method uses probabilistic weights, we need to use the same type of classifier for the environmental and wearable components in order to guarantee the mathematical significance of the fusion.

Finally, the manual segmentation performs about 6% better than the automatic segmentation. In order to understand which segmentation approach should be chosen in an interaction system, the trade-off between the cognitive load on the user and the effect of a lower accuracy of the gesture recognition should be evaluated.

5.2 Interaction Opportunities and Limitations

The quantitative evaluation on our dataset shows good performances in terms of accuracy and F_1-score. However, our prototype allowed us to investigate also more qualitative features of a synergistic paradigm in the context of the interaction in a car. In Section 3.1 we introduced eight design parameters; in this section, we present how they influence the implementation of our opportunistic system.

A direct consequence of the wearable paradigm is that the **interaction area** is no more limited to some spots in the car but can be extended to the whole car. Once the communication between the wearable and the environmental systems is established, it is possible to control the IVIS everywhere in the car.

Gestures can be user dependent. **Personalized** configurations and profiles can be shared from the wearable system to the environmental one. In addition, a wearable system can be designed to share the user information with an unknown system. For example, in a car sharing scenario, the wearable information of a user can be used to configure a new car automatically: binding the personalized gesture vocabulary to the system controls.

With a synergistic system an easy way to increase **consistency** is to online adapt the weights used in the fusion module to achieve better performances. For example, it is possible to penalize or award a classifier directly decreasing or increasing the weights used by the SR or the MCC methods.

The car is an interaction milieu that is generally considered as private. Therefore, in this specific context the needs of **private interaction and intimate interfaces** are reduced, even though the wearable components represent a good solution for this problematic.

The environmental component of a synergistic system can provide **localized information** to the driver taking into account the specificities of the car. For example, the windshield can be used to display information to the user that does not need to move the gaze from the road conserving the safety of the driving. As mentioned before, we placed the 5 pressure sensors in specific regions of the external ring of the steering wheel to have **localized controls** helping the driver to keep the hands on the steering wheel while interacting with the IVIS.

The environmental sensing can be directly integrated in the car exploiting the existing processing and power resources. A synergistic paradigm allows the wearable system to take advantage of the vehicle **resources availability**. In fact, even if the wearable components still require energy to sense the information, the processing can be deployed at environmental side. The downside of this approach is that the transmission of the information can be highly demanding in term of energy and may slow down the whole processing system. Therefore, accurate analyses should be performed case-by-case.

The presence of different gesture lexicons for the wearable and the environmental paradigm can be treated as **resource management**. Commands linked to secondary tasks can be integrated as wearable components and made available for the driver as well as the passengers of the vehicle; on the other hand, primary commands, that can affect passengers' safety, can be made accessible only to the driver by localizing the control in a particular spot of the vehicle.

The previous examples can be used as guideline to tune the design parameters of an in-vehicle opportunistic interaction system in order to take advantage of the strong features of both the environmental and the wearable systems.

6. CONCLUSION

In this paper, we presented an opportunistic system for natural interaction in a vehicle. We proposed a synergistic approach combining wearable and environmental paradigms to recognize micro-gestures on the steering wheel. We compared different

Table 1. Fusion results (μ and σ) F1 score and accuracy. In red the best single classifier; in bold the best fusion methods.

(μ|σ)	Manual Segment. OR		Automatic Segment. OR		Manual Segment. AND		Automatic Segment. AND		Manual Segment. ADAPTIVE		Automatic Segment. ADAPTIVE	
Method	**F1**	**A**	**F1**	**A**	**F1**	**A**	**F1**	**A**	**F1**	**A**	**F1**	**A**
Pressure	0.72|0.16	0.79|0.14	0.66|0.20	0.76|0.17	0.72|0.17	0.76|0.15	0.65|0.15	0.76|0.13	0.72|0.16	0.76|0.13	0.65|0.16	0.76|0.12
EMG	0.85|0.12	0.86|0.13	0.73|0.19	0.75|0.19	0.85|0.12	0.85|0.13	0.75|0.19	0.74|0.19	0.85|0.11	0.85|0.12	0.74|0.19	0.72|0.19
SR	0.79|0.13	0.85|0.11	0.70|0.15	0.84|0.09	0.77|0.18	0.72|0.24	**0.78|0.12**	**0.73|0.14**	0.81|0.12	0.82|0.12	**0.74|0.11**	**0.77|0.11**
NB	0.79|0.14	0.84|0.10	0.71|0.15	0.84|0.08	0.74|0.21	0.70|0.26	0.76|0.12	0.70|0.14	0.78|0.11	0.79|0.11	0.71|0.16	0.74|0.15
NBW	0.80|0.14	0.85|0.10	0.71|0.15	0.84|0.08	0.75|0.21	0.71|0.27	0.77|0.13	0.71|0.14	0.80|0.12	0.80|0.12	0.71|0.15	0.74|0.15
MCC	0.81|0.14	0.87|0.09	**0.72|0.14**	**0.86|0.09**	0.80|0.20	0.75|0.26	**0.78|0.13**	**0.73|0.15**	0.84|0.11	0.85|0.10	**0.74|0.13**	**0.77|0.12**
MCC+	0.81|0.15	0.87|0.10	0.69|0.16	0.83|0.14	0.78|0.20	0.73|0.25	**0.78|0.16**	0.72|0.17	0.83|0.12	0.85|0.11	0.73|0.16	0.76|0.15
SR*	0.79|0.13	0.85|0.11	0.70|0.15	0.84|0.09	0.77|0.18	0.72|0.24	**0.78|0.12**	**0.73|0.14**	0.81|0.12	0.82|0.12	**0.74|0.11**	**0.77|0.11**
NB*	0.79|0.14	0.84|0.10	0.71|0.15	0.84|0.08	0.74|0.21	0.70|0.26	0.76|0.12	0.70|0.14	0.78|0.11	0.79|0.11	0.71|0.16	0.74|0.15
NBW*	0.80|0.14	0.85|0.10	0.71|0.15	0.85|0.08	0.75|0.21	0.71|0.27	0.77|0.13	0.71|0.14	0.80|0.12	0.80|0.12	0.71|0.15	0.74|0.15
MCC*	**0.83|0.15**	**0.89|0.10**	**0.72|0.14**	**0.86|0.09**	0.80|0.21	0.75|0.26	**0.78|0.15**	**0.73|0.16**	**0.85|0.13**	**0.86|0.12**	**0.74|0.14**	**0.77|0.13**
MCC+*	0.82|0.15	0.88|0.10	0.71|0.18	0.84|0.15	**0.81|0.21**	**0.76|0.27**	0.77|0.17	0.72|0.18	**0.85|0.13**	**0.86|0.12**	0.72|0.16	0.76|0.15

fusion methods and segmentation strategies. The proposed ADAPTIVE segmentation strategy and the MCC* fusion method showed the best performance. In fact, we observed that the synergistic approach can perform equal or better than the best stand-alone classifier.

The proposed system requires an accurate design to satisfy the requirements of synergy. We presented eight parameters to be used for tuning and improving the interaction design by taking into account the opportunities and limitations of the wearable and environmental paradigms. We applied the proposed opportunistic paradigm to the in-vehicle interaction scenario; however, such concept can be easily extended in other areas of HCI.

We observed that the automatic segmentation has a recognition rate of about 6% lower than the manual segmentation. Further analyses, in order to decrease this gap, are planned as future work. Moreover, usability tests will be conducted to understand the impact of the segmentation strategy on the car driver in driving simulation scenario. The adoption of an eye-tracking system will allow us to measure the distraction of the driver and compare it with a standard approach, e.g., buttons or touch screen on the dashboard. Further observations will aim at analyzing the effects of the proposed approach on the driver's behavior.

7. ACKNOWLEDGMENTS

This work has been supported by Hasler Foundation in the framework of "Living in Smart Environments" project.

8. REFERENCES

[1] Riener, A. Driver-vehicle confluence or how to control your car in future? In *Proc. of AutomotiveUI '12*, no. c, p. 217, 2012.

[2] Riener, A. Gestural Interaction in Vehicular Applications. *Computer*, vol. 45, no. 4, pp. 42–47, Apr. 2012.

[3] Rhodes, B. J., Minar, N., and J. Weaver, J. Wearable computing meets ubiquitous computing: reaping the best of both worlds. In *Digest of Papers International Symposium on Wearable Computers*, pp. 141–149.

[4] Wolf, K., Naumann, A., and Rohs, M. A Taxonomy of Microinteractions: Defining Micro-gestures based on Ergonomic and Scenario-dependent Requirements. In *Proc. of the 13th IFIP INTERACT*, 559–575, 2011.

[5] Pfleging, B., Schmidt, A., Döring, T., and Knobel, M. AutoNUI. In *Proc.of AutomotiveUI '11*, 2011, p. 207.

[6] González, I. E., Wobbrock, J. O., Chau, D. H., Faulring, A., and Myers, B. A. Eyes on the road, hands on the wheel: thumb-based interaction techniques for input on steering wheels. In *Proc. of Graphics Interface 2007*, 95–102, 2007.

[7] Döring, T., Kern, D., Marshall, P., Pfeiffer, M., Schöning, J., Gruhn, V., Schmidt, A. Gestural interaction on the steering wheel: reducing the visual demand. In *Proc. of CHI '11*, pp. 483–492, 2011.

[8] Mahr, A., Endres, C., Müller, C., and Schneeberger, T. Determining human-centered parameters of ergonomic micro-gesture interaction for drivers using the theater approach. In *Proc. of AutomotiveUI '11*, 2011, p. 151.

[9] Angelini, L., Caon, M., Carrino, F., Carrino, S., Lalanne, D., Abou Khaled, O., and Mugellini, E. WheelSense: Enabling Tangible Gestures on the Steering Wheel for In-Car Natural Interaction. In *Proc. of HCII 2013* LNCS Vol. 8005, 2013, pp 531-540.

[10] Khushaba, R. N., Kodagoda, S., Liu, D., and Dissanayake, G. Muscle computer interfaces for driver distraction reduction. *Computer methods and programs in biomedicine*, vol. 110, no. 2, pp. 137–49, May 2013.

[11] Carrino, F., Carrino, S., Caon, M., Angelini, L., Abou Khaled, O., and Mugellini, E. In-Vehicle Natural Interaction Based on Electromyography. In *Proc. of AutomotiveUI '12*, pp. 1–3, 2012.

[12] Bose, R., Brakensiek, J., and Park, K. Terminal mode. In *Proc. of AutomotiveUI '10*, 2010, p. 148.

[13] Carrino, S., Mugellini, E., Abou Khaled, O., Ingold, R. ARAMIS: Toward a Hybrid Approach for Human-Environment Interaction. In *Proc. of HCII 2011,* Vol. 6763, 2011, pp 165-174.

[14] Wachs, J. P., Kölsch, M., Stern, H., and Edan, Y. Vision-based hand-gesture applications. Communications of the ACM, vol. 54, no. 2, p. 60, Feb. 2011.

[15] Costanza, E., Inverso, S., and Allen, R. Toward subtle intimate interfaces for mobile devices using an EMG controller In *Proc. of CHI '05*, p. 481, 2005.

[16] "*Teaching manual for the formation and examination of driving instructors*", "*Manuel d'enseignement pour la formation et l'examen des moniteurs de conduite*", Département fédéral de justice et police : http://www.asa.ch/media/archive1/Shop/Ausbildungsunterlagen/gratis/Leitfaden_Fahrlehrer_Fachgruppe_7_f.pdf, (June 2013).

[17] Sokolova, M., and Lapalme, G. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

[18] Kittler, J., and Alkoot, F. M. Sum versus vote fusion in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, pp. 110–115, Jan. 2003.

[19] Kittler, J., Hatef, M., Duin, R. P. W., and Matas, J. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, Mar. 1998.

[20] Kuncheva, L. I., Bezdek, J. C., and Duin, R. P. W. Decision templates for multiple classifier fusion: an experimental comparison. *Pattern Recognition*, vol. 34, no. 2, pp. 299–314, Feb. 2001.

[21] Matthews, B. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 1975.

[22] Jurman, G., Riccadonna, S., and Furlanello, C. A comparison of MCC and CEN error measures in multi-class prediction. *PloS one*, vol. 7, no. 8, p. e41882, Jan. 2012.

[23] Caputo, B., Sim, K., Furesjo, F., and Smola, A. Appearance-based object recognition using SVMs: which kernel should I use? *Proc. of Neural Information Processing Systems Workshop on Statistical methods for Computational Experiments In Visual Processing and Computer Vision*. Whistler, 2001.

[24] Carrino, S., Péclat, A., Mugellini, E., Abou Khaled, O. and Ingold, R. Humans and smart environments: a novel multimodal interaction approach. *Proc. ICMI '11*, 2011, p. 105.

[25] Souza, C. http://sourceforge.net/projects/accord-net/, 2013. (June 2013).

[26] Tekscan FlexiForce: http://www.tekscan.com/flexible-force-sensors, (June 2013).

[27] Cram, J. 2011. *Cram's introduction to surface electromyography*. Jones & Bartlett Publishers.