Exploring Interaction Concepts for a Context-aware Smart Office Prototype

Master's Thesis in Medical Engineering

submitted
by

Robert Richer

born 04.06.1994 in Bamberg

Written at

Pattern Recognition Lab, Lehrstuhl für Informatik 5 (Mustererkennung)
Department for Computer Science
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

in Cooperation with

Responsive Environments Group, MIT Media Lab
Massachusetts Institute of Technology

Advisors:
Dipl.-Ing. Nan Zhao (Responsive Environments Group, MIT Media Lab)
Prof. Joseph A. Paradiso, PhD (Responsive Environments Group, MIT Media Lab)
Prof. Dr. Matthias Berking (Chair of Clinical Psychology and Psychotherapy)
Prof. Dr. Bjoern Eskofier (Digital Sports Group, Pattern Recognition Lab)

Started: 01.12.2016

Finished: 31.05.2017

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Richtlinien des Lehrstuhls für Studien- und Diplomarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts.

Erlangen, den 6. Juli 2017

Übersicht

Das aufstrebende Potential neuer Interaktionskonzepte offenbart viele Möglichkeiten für adaptive Umgebungen. Da die Umgebung immer mehr vernetzt wird und abhängig vom aktuellen Kontext reagiert, wird es auch komplexer, die Umgebung zu steuern, was den aktuell etablierten Ansatz, Kontrolldimensionen manuell anzupassen, schwierig und umständlich macht. Im Gegensatz dazu können andere Modalitäten, wie *Smart Agents* und *Brain Computer Interfaces* diese Herausforderungen bewältigen. Aus diesem Grund befasst sich diese Arbeit mit der Analyse dieser verschiedenen Interaktionsmöglichkeiten, die als eine Art Vermittler zwischen dem Menschen und seiner Umgebung fungieren sollen. Verschiedene Smart Agents, die sich hinsichtlich Eingabemodalität und Intelligenz unterscheiden, wurden dafür in einem Prototypen für ein intelligentes Büro, *Mediated Atmospheres*, integriert. In einer Nutzerstudie (N=33) wurde ihre Gebrauchstauglichkeit evaluiert und eroiert, ob sie unterschiedlich von den Teilnehmern wahrgenommen wurden. Das Brain Computer Interface besteht aus einem tragbaren EEG-Stirnband, für das verschiedene Algorithmen zur Echtzeit-Klassifizierung des Mentalzustandes vorgestellt wurden, um zu klassifizieren, ob der Benutzer gerade fokussiert oder entspannt ist. In einer Nutzerstudie (N=11) wurden diese Klassifikatoren evaluiert.

Die Evaluierungsergebnisse der Smart Agents zeigen, dass eine graphische Benutzeroberfläche das am besten bewertete System ist, gefolgt von den Text- und Sprachassistenten. Beide Arten von Assistenten helfen dem Nutzer mit ihrer freundlichen Art allerdings dabei, sich besser im Raum zurechtzufinden. Außerdem konnte ein klarer Unterschied in der Bewertung zwischen Muttersprachlern und nicht Muttersprachlern festgestellt werden, ebenso wie ein Unterschied zwischen den Erwartungen der Benutzer an die Systeme, und der tatsächlichen Erfahrung. Die Evaluierung des Brain Computer Interfaces zeigte, dass der auf der *Tsallis*-Entropie basierende Algorithmus am besten zur Klassifizierung des *Fokus*-Zustandes geeignet ist, wohingegen die *Renyi*-Entropie die besten Ergebnisse bei der Klassifizierung des *Entspannt*-Zustandes erzielt hat. Sensitivitäten von 82.0 % bzw. 80.4 % und Spezifitäten von 82.8 % bzw. 80.8 % wurden bei der Klassifizierung der beiden Zustände erzielt.

Die Ergebnisse dieser Arbeit offenbaren das Potential von Smart Agents für die Interaktion mit intelligenten Umgebungen, z.B. einem Smart Office, um die Interaktion im Arbeitsalltag zu erleichtern. Darüber hinaus zeigt es die Möglichkeit auf, ein tragbares EEG-System für eine Erkennung des Mentalzustandes in Echtzeit einzusetzen, und es zur Steuerung von intelligenten Umgebungen, wie z.B. *Mediated Atmospheres*, zu verwenden.

Abstract

The emerging potential of new interaction concepts unveils new possibilities for adaptive environments. By becoming more pervasive and context-aware, the control complexity increases likewise, making the traditional approach of the user manually adjusting those dimensions cumbersome and tedious. Other modalities like conversational agents and brain computer interfaces could cope better with those challenges. Therefore, this work explores the application of those new interaction concepts as a mediator between occupant and environment. Different agents were applied as personal assistants for *Mediated Atmospheres*, an adaptive smart office prototype, each of them differing in their level of system intelligence and input modality. Through a user study (N=33), their usability was evaluated as well as analyzed, whether the agents create a different perception on the user. The brain computer interface was implemented using a wearable EEG headband. Therefore, different measures (one naive measure and four entropy-based measures) were presented for real-time mental state recognition and to quantify the occupants' current levels of being focused and relaxed. Classifiers for Focus and Relax state detection, based on the estimation of probability distributions for the different measures, were developed and evaluated in a user study (N=11).

Results for the agent evaluation show that a graphical user interface as the most familiar system is also the most favorable system, followed by conversational text and voice agents, that both help the user fulfilling the required tasks in an engaging and friendly way. Furthermore, a clear difference in perception between native and non-native speaker could be observed, as well as a disparity between the users' expectations and experience. The evaluation of the brain computer interface shows that the measure based on *Tsallis* entropy performed best for the Focus score, whereas the *Renyi* measure performed best for the Relax score. Sensitivities of 82.0 % and 80.4 %, with specificities of 82.8 % and 80.8 %, were achieved for the Focus and Relax scores, respectively.

The findings reveal the potential of conversational agents for the interaction with adaptive environments such as a smart office to reduce work and information overload. Furthermore, it demonstrated the possibilities of using a wearable EEG system for real-time mental state recognition and to control adaptive, context-aware environments like *Mediated Atmospheres*.

Contents

Chapter 1

Introduction

Smartphones have become an essential part of our daily life and are deeply embedded in our everyday actions, in a personal as well as in a professional environment [Smi15]. According to [And15], smartphones are used about 85 times a day, with more than five hours spent in total with the device. Because interacting with a smartphone mostly requires to use one or both hands and hence easily distracts from other activities [Smi15], companies like Apple (*Siri*, 2011), Google (*Google Now*, 2012), Microsoft (*Cortana*, 2015), and Amazon (*Alexa*, 2015) have released smart voice assistants that are integrated into smartphones, computers, or Internet of Things (IoT) devices and enable a hands-free and more unobtrusive interaction. In 2016, Google announced the *Google Assistant* as the successor of *Google Now*. It was designed as intelligent personal assistant that engages the user in a two-way conversation, but moreover offers multimodal input: it is not only possible to talk to the assistant over voice, but also to send text messages using the messaging service Google Allo [Goo16].

Research groups have been developing agents that answer to questions conversationally for the past two decades [Gel12]. However, those companies have made them available for the broad public, following the vision of the HCI community that voice agents will take their place among the most important human computer interaction modalities as they will become more conversational [Bre90].

Similarly, recent developments in ubiquitous computing unveiled a new range of other HCI concepts, such as brain computer interfaces (BCIs) based on measuring the electroencephalogram (EEG), the electrical activity of the brain using low cost wearable EEG devices. For instance, recent work used wireless EEG devices for games, or biofeedback applications [Abu16, Amo16, Bas16].

Therefore, the emerging potentials of conversational agents and brain computer interfaces offer new possibilities for the interaction with advanced context-aware applications, especially smart environments. Equipped with pervasive and ubiquitous technologies, those environments obtain information about itself and the actions and preferences of the user, while concurrently increasing the complexity of control dimensions. By sensing the users' requests and responding to them, complex, seemingly intelligent tasks may be performed automatically [Kid99], creating a symbiosis between the space and the occupant for providing meaningful interactions with his or her surroundings.

For this work, *Mediated Atmospheres*, proposed by Zhao et al. [Zha17] was used as one scenario of an intelligent, adaptive environment. It aims to dynamically influence the users' abilities to focus on a specific task or to recover from a stressful situation. Multimodal media – lighting, video projection, and sound – is used to seamlessly transform the appearance of the workspace. Different environments can be projected into the room according to the occupants' preferences or different work scenarios. In order to unleash the full potential of such an adaptive environment, the user experience should not be compromised by a cumbersome and tedious interaction with the workspace. Therefore, the traditional approach of a discrete number of control dimensions that have to be manually adjusted by the user is obsolete. The vision should rather be a system that acts as a mediator between person and environment, reducing the work and information overload [Mae94], and being capable of understanding requests and seamlessly translating them into a manifestation of lighting, video and sound.

Nowadays, conversational agents still suffer from a "gulf of evaluation" between the users' expectations and the system operation [Lug16]. Naturally, users have higher expectations towards a system that pretends to be capable of conducting a conversation, and hence tend to be more disappointed if the system provokes wrong actions [Lug16]. However, it can be assumed that the "gulf of evaluation" depends on the input modality, and is wider for voice-based input than for text-based input.

Therefore, this work explores the application of different agents as personal assistants for multimodal mediated environments using the *Mediated Atmospheres* framework. As each interface has its strengths and weaknesses, and is applicable for different scenarios, depending on the current task and the user preference, four different systems are presented and compared. Each of them are integrated into the office environment, intended to increase both the system usability and the user experience by providing a natural way of interaction. The best system for the interaction with the smart workspace is expected to be found by creating and evaluating agents that differ in their level of system intelligence (i.e. context-awareness, and variety of features) and input modality (i.e. voice-based or text-based input). Additionally, the application of brain computer interfaces using a wearable EEG headband is explored. Information from this device is used to perform a real-time mental state recognition of the occupant in order to determine the current levels of Focus and Relax and to transform the appearance of the office space accordingly. Because this problem requires fast and robust algorithms, different approaches (one naive and four entropy-based approaches) are implemented and evaluated in order to determine the most suitable solution.

After relevant related work in Section 2 has been explored and discussed, Section 3 describes the infrastructure of the smart office prototype and the different interaction concepts that were developed within this work. In Section 4, the evaluations, including experiment designs, procedures, and measures, are presented for both smart agents and brain computer interface, with the obtained results shown in Section 5. For the smart agents, subjective measures like the overall usability the different systems, as well as the perception of intelligence, engagement, trust, and control the agents create on users exploring their features are of particular interest. Furthermore, objective measures like fulfillment time, recognition rate, and the number of necessary interactions with a system to fulfill a certain task are defined and recorded. For the brain computer interface, the performance of different algorithms for real-time mental state recognition are compared with regard to parameters like sensitivity, specificity and area under the Receiver Operating Characteristic (ROC) curve. Leveraging the results, the strengths and weaknesses of all approaches are discussed in Section 6, followed by a conclusion in Section 7.

Chapter 2

Related Work

In the past years, researchers have equipped built environments with pervasive and ubiquitous technologies in order to control specific properties like lighting [Ros15, Che13, Ald10, Mag11], glazing regulation [Kal16], HVAC control [Fel10], sound [Kai05], or information display [Ras98, Ben14, Tom08]. Other research groups have created whole adaptive ambient environments to enable life-size telepresence [Pej16], or to transform its appearance and physical shape according to physiological data of the occupant [Sch10]. Another example is the ambientROOM that utilizes subtle cues of sound, light, or motion to provide meaningful information to the user [Ish98].

Similarly, context-aware systems also been explored within the last two decades. According to Abowd et al. [Abo99], context-awareness is defined as "the use of context to provide task-relevant information and/or services to a user", where context can be categorized into *Activity*, *Time*, *Identity*, and *Location* [Per14]. As an example of a context-aware environment, researchers have created the Aware Home, an entire home equipped with ubiquitous computing technologies in order to create a living laboratory that is as realistic as possible [Kid99]. Other examples that aimed to increase convenience are the Neural Network House [Moz98], or the Adaptive House [Moz05].

Additionally, the EEG has been applied by research groups for context-aware applications, such as the assessment and quantification of drivers fatigue in order to increase road safety by using alpha spindles [Sim11] or ratios of EEG spectral components [Jap09]. Additionally, other research groups proposed driver fatigue detection and quantification algorithms using entropy-based measures [Kar10, Zha14]. Steps beyond the assessment of drowsiness and fatigue are the recognition of different mental states based on EEG data [Mül08], as well as the analysis of cognitive performance [Kli99] or attention [Cla02, Kel05]. Many of those experiments used visually evoked potentials that require a distinct visual stimulus in order to segment it from the regular EEG signal. Furthermore, they were conducted under controlled laboratory conditions and with clinical, obtrusive EEG systems. In contrast, low cost EEG headsets have been developed by companies like InteraXon (Toronto, Canada) [Int17], NeuroSky (San Jose, CA, USA) [Neu17], or Emotiv (San Francisco, CA, USA) [Emo17]. Additionally, several patents were published, presenting wearable EEG devices, such as US patent *US12716425* [A.1], US patent *US12955016* [A.2], or US patent *US14216925* [A.3].

In order to evolve from adaptive, context-aware rooms to intelligent environments [Men14], agents have been integrated into the built environment. They are supposed to learn the occupants' preferences and habits and execute appropriate actions automatically. For instance, MavHome is an agent-based smart home with algorithms to predict the inhabitants' actions [Coo03]. Other smart home applications see the benefit of agents in power management [Abr08] or to mediate the interaction with a smart grid [Ala16]. In smart offices, agents have been integrated for a virtual secretary [Dan08] or to accompany and lead persons in an office building [Bag03]. With increasing popularity of smart agents and environments, people have analyzed them in terms of acceptance and trust [Hos13], dominance and cooperativity [Str16], or personality traits [Men16].

The idea of a having an agent, an artificial intelligence designed to maintain a conversation with a human, has already been envisioned by Alan Turing in the 1950s [Tur50] and was successfully implemented for the first time by Joseph Weizenbaum with ELIZA [Wei66]. It is considered the first "chatbot", a text-based conversational agent simulating a natural human interaction between the computer and a user. Systems and methods for conversational agents have also been patented in the past, such as US patent *US8073681* [A.5] and US patent *US7019749* [A.4]. Recently, conversational agents have been created for the touristic domain [D'H15], as a museum guide [Kop05], or in social networks like Facebook or Twitter [Fer16]. The technology of conversational agents has even advanced so rapidly that it has become hard to distinguish between humans and chatbots in social networks [Chu10, Sub16].

Chapter 3

System Description

The present works explores different interaction concepts for a smart office prototype. Therefore, this chapter first introduces the concept of this smart office, *"Mediated Atmospheres"* [Zha17], before the integration of smart agents is described in Section 3.3. In order to determine the best agent for a smart office scenario, different types were implemented. As shown in Figure 3.1, they differ in their input modality (visual vs. voice vs. text) and their system intelligence (basic vs. advanced). All agents communicate with the *Mediated Atmospheres* framework, as visualized as a top-level description in Figure 3.2. Additionally, the integration of a wearable EEG headband as brain computer interface is explored in Section 3.4. Therefore, different algorithms for real-time mental state recognition were developed in this work in order to quantify the levels of Focus and Relax of the occupant.

3.1 Mediated Atmospheres

Mediated Atmospheres is a modular workspace capable of digitally transforming its appearance using controllable lighting, video projection and sound. The physical layout of the office space is sketched in Figure 3.3. 20 individually controllable light fixtures light the walls to the occupants' left and right (wall-washing fixture), and the occupant directly (downlight fixture). It is combined with video projection onto a frosted acrylic display. Without video projection, the display is white and opaque and resembles a wall divider for cubicles. Furthermore, the room is equipped with a four speaker ambisonic sound system. If ambisonic sound data is available, it attempts to reproduce the spatial sound conditions for creating a fully immersive experience.

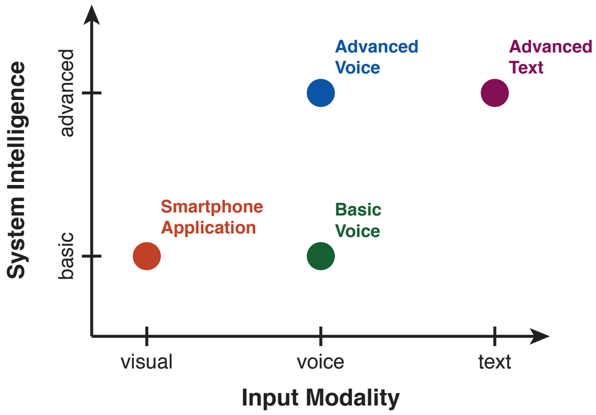


Figure 3.1: Agent Overview. Overview of agents for the interaction with the *Mediated Atmospheres* framework

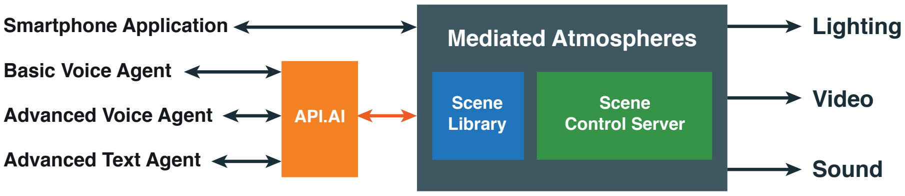


Figure 3.2: Top-level description of system. The interaction modalities communicate with the *Mediated Atmospheres* framework, which maps input commands to control output modalities.

3.1.1 Scene Library

All digital output capabilities are combined to design a library of scenes, multimodal virtual environments, with each of them possibly having a different influence on the occupant's physiology or behavior [Zha17]. Examples of the workspace, transformed into different scenes, are shown in Figure 3.4.

The scene library currently features more than 30 elements, ranging from nature landscapes, like forests, beaches, mountains, over urban scenes, like walking through a city or train rides, to indoor spaces, like libraries or cafés. The scenes were selected to cover a wide range of environments, all creating a different perception on the occupant of the space. Each of the scenes are identified by a unique identifier (Scene ID) and a descriptive name, and contain video data, sound data, and lighting configurations. Furthermore, other different scenes properties were defined, that will later be used by the agents to filter scenes, like:

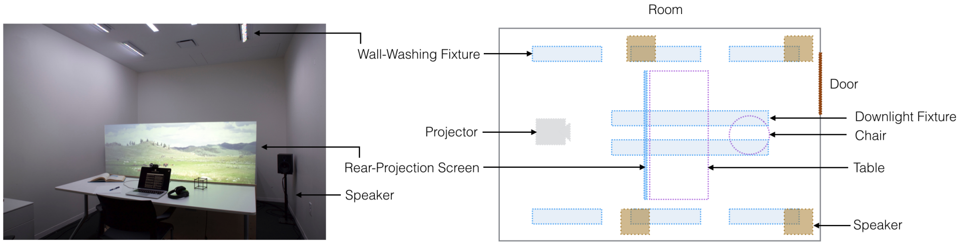


Figure 3.3: Mediated Atmospheres Layout. Physical layout of the office space. [Zha17], with permission.

- *Color Temperature*: Color temperature of the scene's lighting configuration (K)
- *Brightness*: Brightness of the scene's lighting configuration, measured as the horizontal illumination on the desk (lux)
- *Primary Color*: Dominant colors of the scene's video content
- *Light Direction*: Direction of the light scene's lighting configuration, relative to the office desk in the middle of the room
- *Keywords*: Keywords describing the scene's video content. Keywords were generated using *Clarifai*[1], which provides an online API for image and video recognition

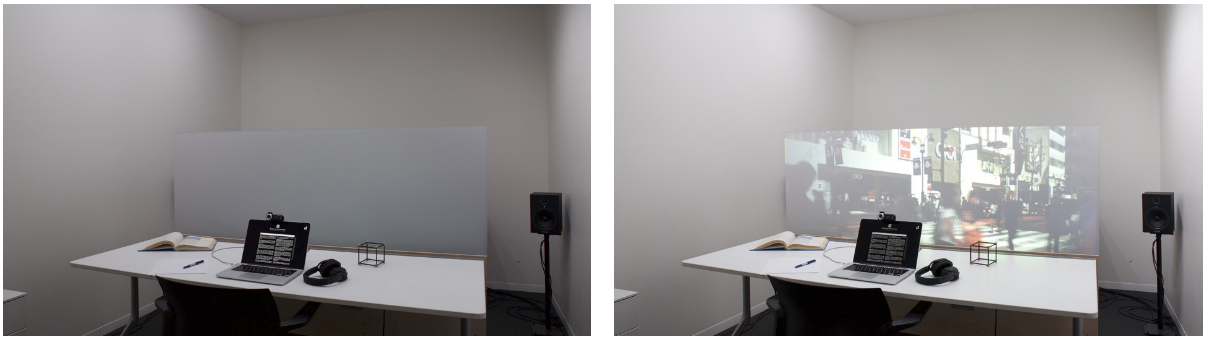
All scene properties were stored in JavaScript Object Notation (JSON). Listing 3.1 shows an example for one scene configuration file.

The scene library was implemented as an Amazon dynamoDB database, a scalable and non-relational database web service, that offers APIs for platforms like Python, Node.js, or Android [Siv12]. It supports multiple indexes, that were utilized to directly scan the database for scenes with certain properties, like a specific color temperature range. Listing 3.2 shows an example in Node.js of how to find a scene based on its descriptive name by performing a query against the scene database, whereas Listing 3.3 shows an example of how to retrieve scenes within a certain color temperature range by performing a scan on the scene database.

Furthermore, every occupant can create his or her own model with scenes that help being more focused or being more relaxed. Therefore, *Focus* and *Relaxation* coordinates are assigned to each scene, and individually stored for every user in the scene library. The model can be created using a website with a graphical user interface, in which scenes

---

[1] http://www.clarifai.com/



(a) Neutral

(b) City

(c) Silverman

(d) Kites

(e) Library

(f) Forest

Figure 3.4: Mediated Atmospheres. The smart office prototype, transforming its appearance into different scenes. [Zha17], with permission.

are placed along a two-dimensional coordinate system, with *Focus* and *Relaxation* being the two axes (see Figure 3.5).

```
{
  "descriptive_name": "Forest",
  "scene_id": 0,
  "feature": {
    "color_temperature": 4600,
    "lighting_metadata": {
      "mean_brightness": 943,
      "number_lights": 2,
      "light_direction": [ "top left", "botom left" ]
    }
  },
  "primary_color": [ "#b39a6d", "#6b6946", "#b3a39e" ],
  "labels": [ "nature", "wood", "trees", "landscape", "outdoors",
        "fall", "river", "no people", "mountains" ],
  "length": "01:00:03",
  "rep_frame_path": "rep_frame/0000_forest.jpg",
  "path": "videos/0000_forest.mp4",
  "lighting": "lighting/0000_forest.json"
}
```

Listing 3.1: Scene Configuration File

3.1.2 Scene Control Server

For controlling the office space in real-time, the *Mediated Atmospheres* framework features a *Scene Control Server*, implemented in Python 3.5. It provides a Websocket Interface in order to allow other applications to communicate with the server using JSON-based messages. Possible commands for the Scene Control Server are listed in Table 3.1. As shown in Figure 3.2, this API is utilized to integrate different applications, that dynamically control the appearance of the workspace.

```
function getScenesByName(sceneName, callback) {
  var docClient = new AWS.DynamoDB.DocumentClient();
  var params = {
    TableName: "SceneLibrary",
    KeyConditionExpression: "#name = :name",
    ExpressionAttributeNames: { '#name': "name" },
    ExpressionAttributeValues: { ':name': sceneName }
  };
  docClient.query(params, function(err, data) {
    if (!err) { callback(data['Items']); }
  });
}
```

Listing 3.2: Query against Scene Library

```
function getScenesByColorTemp(colorTempRange, callback) {
  var docClient = new AWS.DynamoDB.DocumentClient();
  var params = {
    TableName: "SceneLibrary",
    IndexName: "table_color_temperature",
    FilterExpression: "#temp between :temp1 and :temp2",
    ExpressionAttributeNames: { '#temp': "color_temperature" },
    ExpressionAttributeValues:
    { ':temp1': colorTempRange[0], ':temp2': colorTempRange[1] }
  };
  docClient.scan(params, function(err, data) {
    if (!err) { callback(data['Items']); }
  });
}
```

Listing 3.3: Scan on Scene Library

3.1.3 Sensor Collection Server

In order to learn about the occupants' physiology, *Mediated Atmospheres* is equipped with a real-time sensor collection infrastructure implemented in Python. So far, the office prototype supported the *Zephyr Bioharness 3* [Med17] (Medtronic, Dublin, Ireland), a chest strap for physiological monitoring, and *Intraface* [DlT15], a facial feature tracking software for videos captured by a webcam (*Logitech Quickcam Vision Pro*, Apples, Switzerland) on the desk. From the sensors, features like *Heart Rate Variability*, *Respiration Rate*, *Head*

Table 3.1: Scene Control Server Commands

Command Description	JSON Example
On Turns *Mediated Atmospheres* on	`{ type: "ON", name: null , id: -1 }`
Off Turns *Mediated Atmopsheres* off	`{ type: "OFF", name: null , id: -1 }`
Ping Gets the current scene in *Mediated Atmospheres*	`{ type: "PING", name: null , id: -1 }`
Switch Scene Transforms *Mediated Atmospheres* into the specified scene	`{ type: "SCENE", name: scene_name , id: scene_id }`

Orientation, and *Facial Expression* are extracted and used to compute real-time focus and relaxation score of the occupant. As described in Section 3.4, additional features from a *Muse* [Int17] (InteraXon, Toronto, Canada) wearable EEG Headband were added within this work.

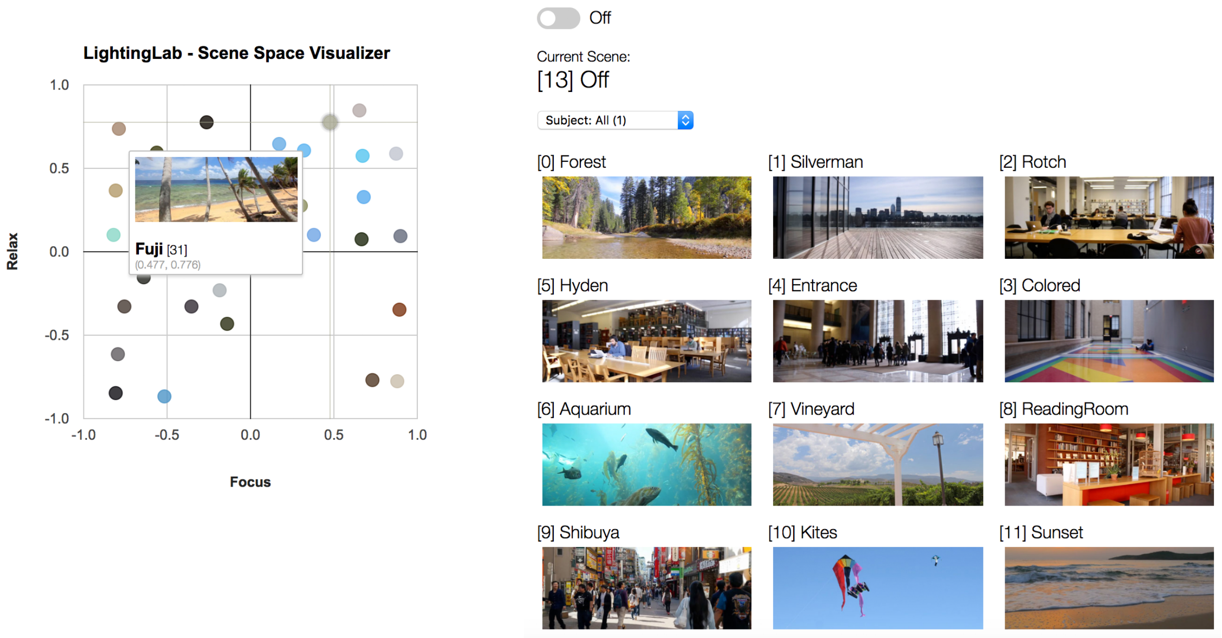
3.2 Smartphone Control

The smartphone application was developed for Android-based mobile devices using the Android SDK 7.0 (API level 24). It communicates with the Scene Control Server using the Websocket API, and is connected to the Scene Library via the Android AWS Mobile SDK for fetching scene information. Therefore, a `DynamoDBWrapper`, initialized with an instance of `AmazonDynamoDBClient`, is used to map a client-side class to the database. For the Scene Library, the mapping class maps each item of the database to a `Scene` Java object that contains each attributes of the table. In order to establish such a mapping, DynamoDB defines annotations to identify table properties, such as table name, table keys, and attributes. Listing C.1 sketches how to create a DynamoDB database client in Android and how to retrieve all scenes by performing a scan on the Scene Library.

The smartphone application has a straightforward and easy to handle WYSIWYG (“What You See Is What You Get”) user interface, allowing users to filter the available

Table 3.2: Feature range of agents. x = available, – = not available.

	Agent			
Feature / Action	*Smartphone*	*Basic Voice*	*Advanced Voice*	*Advanced Text*
Basic Features				
Turn on/off Turns all of the room's output modalities on/off	x	x	x	x
List scenes Lists a subset of available scenes to the user	x	x	x	x
Current scene Tells user the current scene's name	x	x	x	x
Switch scene [name] Changes scene based on a descriptive name	x	x	x	x
Advanced Features				
Switch scene [properties] Changes scene based on their properties (color temperature, brightness, etc.)	x	–	x	x
Switch scene [comparison] Changes scene based on comparison with current scene (warmer, brighter, etc.)	x	–	x	x
Switch scene [keywords] Changes scene based on keywords characterizing the scene content	–	–	x	x
Context-awareness Information about occupant, and about previous interactions	–	–	x	x
Recommend scene [weather] Changes scene based on current weather information	–	–	x	x
Recommend scene [time] Changes scene based on current time of day	–	–	x	x
Recommend scene [mental state] Changes scene based on desired mental state of occupant (focus/relaxed)	–	–	x	x
Visual Features				
Scene Preview Displays a preview image of the scene	x	–	–	–



Current Scene:
[13] Off

Subject: All (1)

Figure 3.5: Scene Model User Interface. Visualization of personalized scene model with *Focus* and *Relax* coordinates assigned to each scene.

scenes by their properties using sliders (see Figure 3.6), see preview images for each scene, and change the environment by selecting a scene. An overview of features available in the smartphone app is listed in Table 3.2.

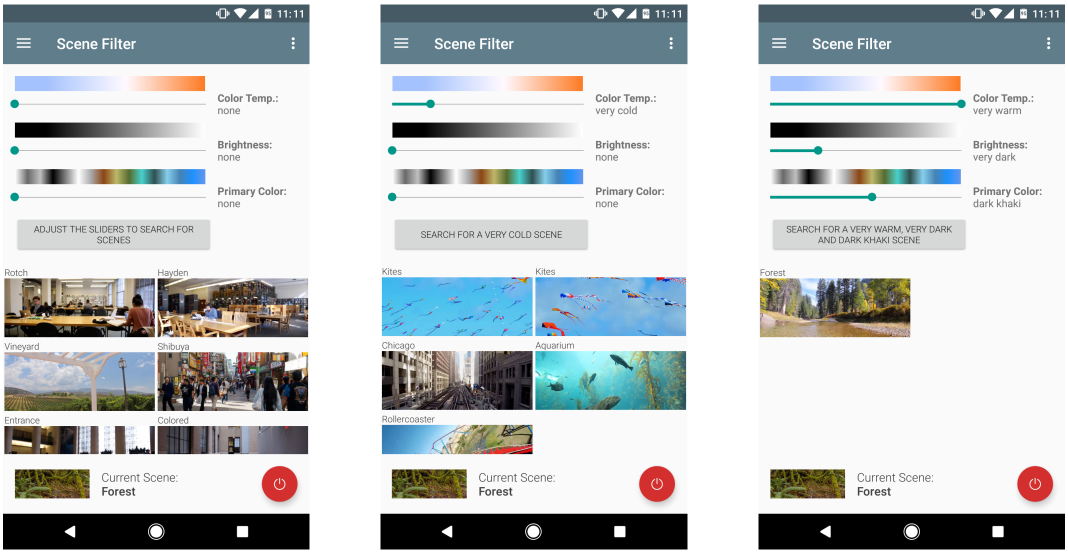
3.3 Smart Agents

All agents were implemented using API.AI [2], a platform for building conversational agents for different applications [API17c]. As sketched in Figure 3.7, it uses *Intents* for representing the mapping between the *User Input* and the *Action* that should be performed by the system. API.AI furthermore provides a fulfillment service which allows to connect the agent to a webhook, that is being executed if an intent was triggered. In this work, a Node.js based webhook implementation was used to perform further logic, to create more diversified agent responses, to connect to external APIs, and to communicate with *Mediated Atmospheres*. The agents were first developed and tested locally using *ngrok* [3] which creates a secure HTTPS connection to a local webserver in order to fulfill webhooks

---

[2] https://api.ai

[3] https://ngrok.com



Color Temp.: none
Brightness: none
Primary Color: none
Current Scene: Forest

Color Temp.: very cold
Brightness: none
Primary Color: none
Current Scene: Forest

Color Temp.: very warm
Brightness: very dark
Primary Color: dark khaki
Current Scene: Forest

(a) No filter **(b)** Very cold scenes **(c)** Very warm, very dark and khaki scenes

Figure 3.6: User interface of the Smartphone Application. Adjustable sliders to filter scenes for *Mediated Atmospheres*

locally. The final version was then deployed to the *Google Cloud Platform* [4].

The agent designed in API.AI can be integrated into different platforms, such as *Actions on Google*, *Facebook Messenger*, *Slack*, *Twitter*, *Skype*, or *Microsoft Cortana* [API17b]. For the voice agents, *Actions on Google* was used. Therefore, API.AI wraps the functionality of the *Google Actions SDK* and provides additional features, such as a graphical user interface, natural language understanding, and machine learning [oG17a]. Integrating the agent into Actions on Google allows to deploy it onto Google Home, a wireless smart speaker running the *Google Assistant*, an intelligent personal assistant designed to engage the user in a two-way conversation [Pic16]. Because the Google Assistant runs one built-in and multiple third-party agents, the user has to invoke one particular agent either by *Name Invocation*, like *"Okay Google, let me talk to {agent_ name}"*, or by *Deep Link Invocation*, like *"Okay Google, tell {agent_ name} to {specific_ action}"*. To facilitate a differentiation between the built-in actions and the actions of the *Mediated Atmospheres* agents, a different female voice was used for the agents' text-to-speech engine, recommended for news, education, and game applications [oG17b].

The text agent was integrated as a chat bot into the Facebook Messenger platform [Ros16]. The API.AI agent is connected to a Facebook App hosted by a Facebook

---

[4] https://cloud.google.com

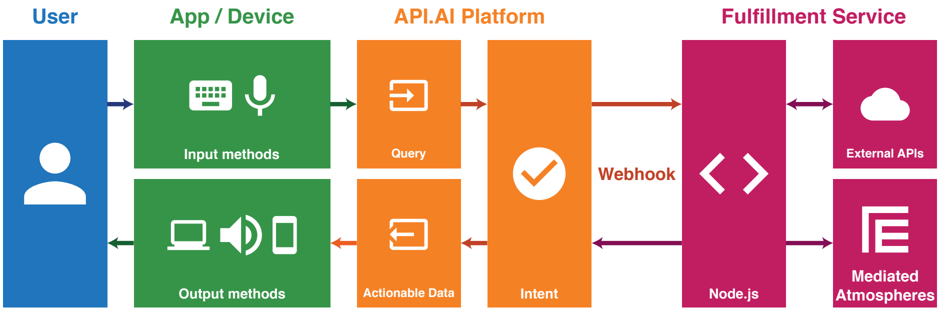


Figure 3.7: Top-level description of agent implementation. The API.AI platform processes user input, maps it to intents and performs corresponding actions. Via Webhook, it is connected to external services and *Mediated Atmospheres*.

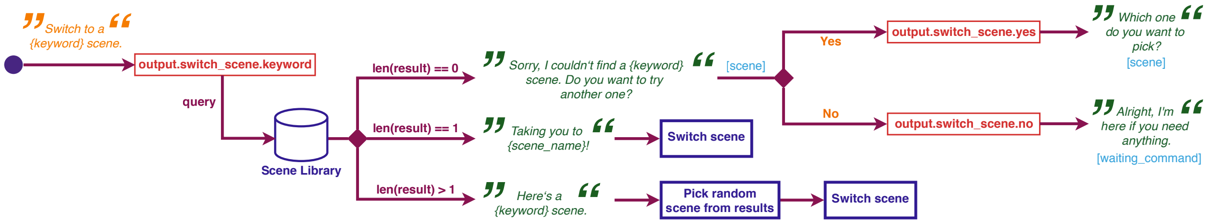
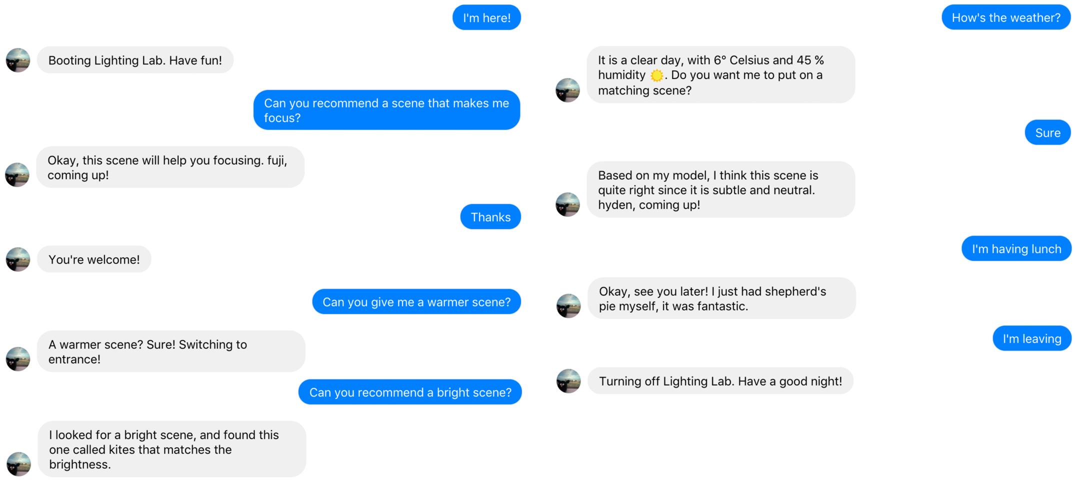


Figure 3.8: Flow chart of scene switching dialog. The dialog is triggered when the user wants to switch to a scene based on a keyword. *Red – Invoked action; Purple – System function; Blue – Current context; Orange – User Input; Green – Agent response.*

Page. This allows the user to directly text the agent with the desired action, without any additional invocation. In order to make the agent more engaging, emojis were included into its text responses.

3.3.1 Basic Agent

This voice agent was designed to cover the basic requirements of an agent for the smart office prototype, like turning the room on/off, and switching scenes. The whole list of features is listed in Table 3.2. The *Basic Voice Agent* is not context-aware, which means that it has no information about the users' preferences, or if the current user has been in the office before. Therefore, it always offers to give an introduction about the features every time an action was invoked. The basic agent is non-conversational, meaning that the interaction with the user is based on a simple *Request – Action* scheme. Therefore, it ends the conversation after successfully receiving a command rather than engaging the



I'm here!

Booting Lighting Lab. Have fun!

Can you recommend a scene that makes me focus?

Okay, this scene will help you focusing. fuji, coming up!

Thanks

You're welcome!

Can you give me a warmer scene?

A warmer scene? Sure! Switching to entrance!

Can you recommend a bright scene?

I looked for a bright scene, and found this one called kites that matches the brightness.

How's the weather?

It is a clear day, with 6° Celsius and 45 % humidity ☀. Do you want me to put on a matching scene?

Sure

Based on my model, I think this scene is quite right since it is subtle and neutral. hyden, coming up!

I'm having lunch

Okay, see you later! I just had shepherd's pie myself, it was fantastic.

I'm leaving

Turning off Lighting Lab. Have a good night!

Figure 3.9: Example dialog of *Advanced Text Agent*. The text agent is embedded into Facebook Messenger and can be accessed via website or Facebook Messenger app.

user in a dialog. As the agent has no information about scenes, users have to directly tell the agent which scene to switch to using the scene's descriptive name, either by having prior knowledge about certain scenes, or by asking the agent to list a subset of scenes.

3.3.2 Advanced Voice and Text Agents

The advanced agents exceed the basic agent in terms of both feature range (see Table 3.2), and context-awareness. Figures 3.8 shows an example flow chart between user and agent and possible responses, depending on the users' input. Further flow charts are sketched in C.2. The advanced agents are implemented as conversational agents and incarnate a personal assistant for the smart office. Therefore, they are talking about themselves in first person and are referring to the occupants by their name. This allows the agents to remember previous interactions with users and hence skip the introduction procedure if they are already familiar with the system. As shown in an example dialog in Figure 3.9, both advanced agents offer a wider range of responses to hold the conversation, depending on the current context. For example, different responses are provided if users are leaving the room at the end of the day, or just to have a break, like having lunch.

Because both advanced agents have information about the properties of scenes, users can not only switch to a scene by specifying its descriptive name, but are also able to select scenes by more abstract commands, like "Can you switch to a city/nature scene?", or "Can you switch to a bright scene?". It is further possible to to combine scene properties, like

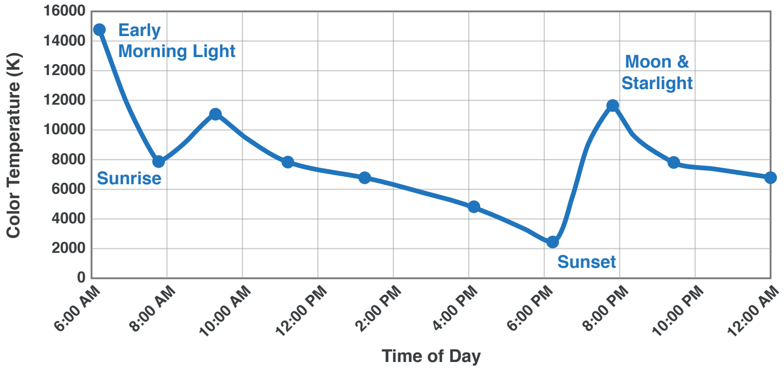


Figure 3.10: Color temperature of natural light. Depending on the current time of day, the color temperature of natural light can range between $2,000\ K$ and $15,000\ K$ [Lum13].

"Take me to a warm and bright outdoors scene", or to switch to a scene with properties relative to the current scene, like "Switch to a warmer/colder/brighter/darker scene".

However, these features are also more or less present in the smartphone application, and still require the user to map his or her scene expectation to the different properties. For that reason, the agents are also capable of recommending scenes based on the current context, or the user's preferences, such as:

- *Weather*: Recommending a scene according to the current weather is based on the mapping of real-time weather information to scene properties as described in Table 3.3. The general weather situation was mapped to the scene's color temperature to approximate the current lighting situation outside. Because high levels of humidity were found to be a key factor for concentration loss and sleepiness [How84], the relative air humidity was linearly mapped to the brightness, with high levels of humidity resulting in scenes with high brightness.

- *Time*: Recommending a scene according to the current time is based on the change of the natural light's color temperature over the day (see Figure 3.10) [Lum13]. Therefore, warmer scenes are recommended in the afternoon and evening, and scenes with colder or neutral color temperature in the morning and noon in order to match the human circadian rhythm [Duf09].

- *Desired mental state*: The advanced agents are aware of the current occupant, allowing them to recommend scenes to based on his or her preferences, like scenes that help users being more focused or relaxed. As described in Section 3.1.1, users can create their own models that can be then accessed by the agent.

Table 3.3: Mapping of weather information and scene properties

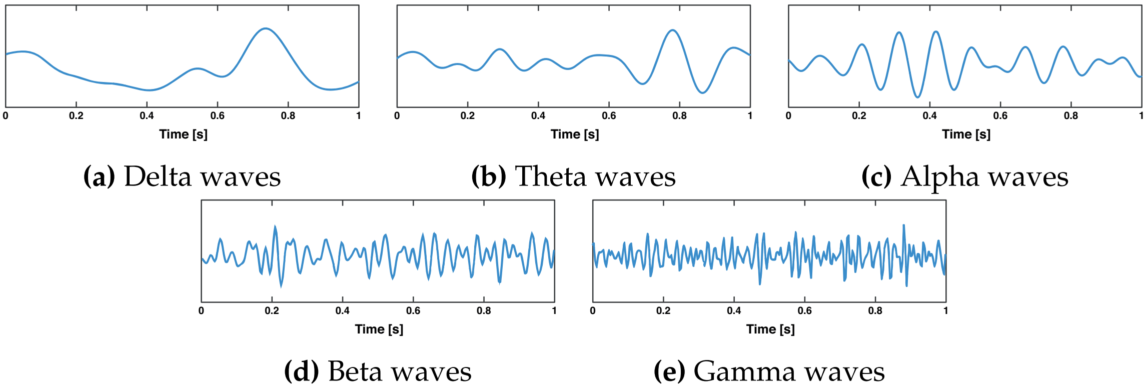
Weather information	**Scene property**
Weather summary:	Color temperature:
clear	warm
(partly) cloudy	neutral
rain, snow, fog	cold
Relative humidity	Brightness
Temperature:	Keywords:
> 25 °C	summer
< 0 °C	winter

3.4 Brain Computer Interface

3.4.1 EEG Basics

The electroencephalogram (EEG) is a noninvasive method for monitoring and analyzing the brain activity. Traditionally, EEG is a standard method in neuroscience and cognitive science, with clinical applications for sleep and memory research, epilepsy monitoring, or attention deficit hyperactivity disorder (ADHD) [Tha13]. It is usually acquired with electrodes placed along the scalp and measures the activity of the brain as oscillations of electrical potential in the cortex. The electrodes are placed according to the 10-20 system, a standardized method for describing and applying electrodes on the scalp [Kle99].

The EEG consists of a variety of frequencies that are associated with different mental states and are traditionally divided into five frequency bands [Cla98]. Table 3.4 shows the division used by the Muse headband with its corresponding frequency ranges. Example signals for each frequency band are visualized in Figure 3.11. Alpha waves occur in the EEG of almost all healthy persons when they are awake but in a quiet, resting mental state and when their eyes are closed, whereas theta waves can be observed in the EEG of adults during phases of drowsiness and hypnagogia. As soon as the person focuses attention to a specific task or mental activity, synchronous alpha waves are replaced by beta and gamma waves of higher frequency. Beta waves occur when performing basic cognitive tasks, whereas gamma waves especially arise during phases of high level mental processing like binding of senses or being focused. The EEG signal is additionally overlaid by both high frequency and low frequency artifacts. Whereas high frequency noise occurs due to thermal noise and power frequency noise (at 50 Hz or



(a) Delta waves **(b)** Theta waves **(c)** Alpha waves

(d) Beta waves **(e)** Gamma waves

Figure 3.11: EEG waveforms. Example signals with durations of 1 s for each frequency band.

Table 3.4: EEG frequency bands provided by Muse headband [Int17].

Name	**Frequency range**	**Associated Mental State** [Tha13, Row85, Sch77]
δ (Delta)	1 - 4 Hz	Deep sleep
θ (Theta)	4 - 8 Hz	Drowsiness, hypnagogia
α (Alpha)	7.5 - 13 Hz	Awake, but quiet and relaxed
β (Beta)	13 - 30 Hz	Basic cognitive tasks
γ (Gamma)	30 - 44 Hz	High levels of mental processing, binding of senses

60 Hz, respectively), low frequency noise mainly origins from eye movement and muscle artifacts [Tha13].

3.4.2 Data Acquisition

The EEG data is recorded using a Muse Headband (InteraXon Inc., Toronto, Canada) [Int17] [5]. It is a commercially available and portable EEG system with four active electrodes (denoted as channels 1-4) located at 10-20 positions *TP9*, *AF7*, *AF8*, and *TP10*, and a common mode reference electrode at *Fpz*, which also acts as driven right leg (see Figure 3.12).

The headband initially oversamples the EEG signal at a sampling rate of 12 kHz and subsequently downsamples it to yield an output sampling rate of 220 Hz. An on-board digital signal processor (DSP) performs noise filtering operations and applies Fast Fourier

---

[5] http://www.choosemuse.com

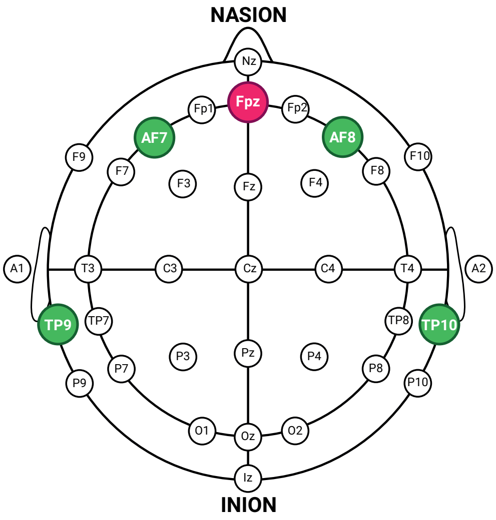


Figure 3.12: Electrode positioning of Muse headband. Electrodes are placed according to the 10-20 system. *Green:* Channel electrodes; *Red:* Reference electrode.

Transform (FFT) on the raw signal using a 256-sample window with an overlap of 90 %, resulting in an output sampling rate of 10 Hz [Int17]. Relative frequency band powers are computed in hardware from raw FFT values as percentages of linear-scale band powers in each frequency band (see Table 3.4).

The acquired data are streamed to a computer running an instance of the *MuseIO* application included in the SDK [6]. It handles the communication with the Muse headband and passes EEG data, formatted as Open Sound Control (OSC) messages, to a client via User Datagram Protocol (UDP). In this work, a *Score Processor*, implemented in Python, acts as client and receives the data from the *MuseIO* application for further data processing. The final score values are transmitted to the *Sensor Collection Server* of *Mediated Atmospheres* via OSC messages. Additionally, a smartphone application was developed for both data acquisition and processing, and the visualization of the raw EEG signals as well as the final *Focus* and *Relax* scores. The smartphone application also transmitted the score values to the *Sensor Collection Server*, as sketched in Figure 3.13.

3.4.3 Data Processing

Live data received from the Muse headband were first preprocessed before different methods of computing Focus and Relax scores (a *naive* as well as *entropy*-based measures) were performed.

---

[6] http://developer.choosemuse.com/research-tools/museio

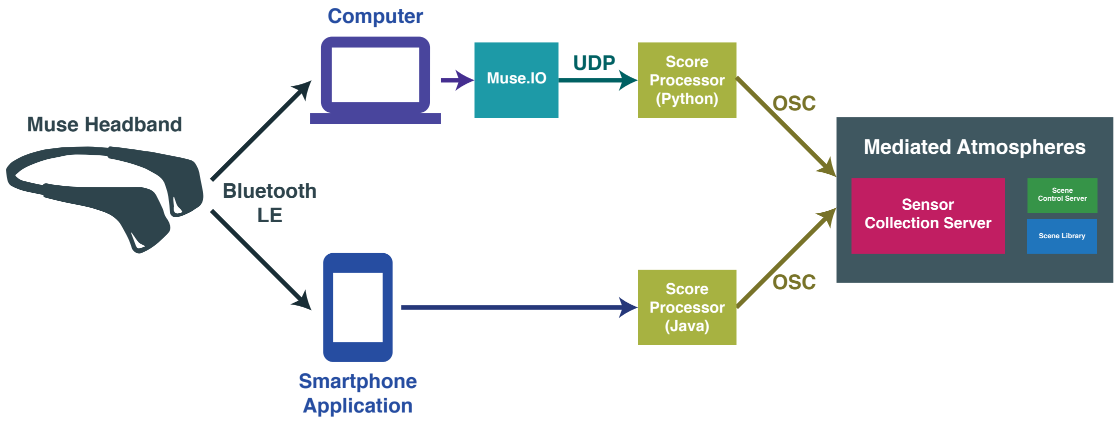


Figure 3.13: EEG Data Acquisition Pipeline. The raw signals from the Muse headband can either be streamed to a computer or a smartphone application. The computed score values are then transmitted to the *Mediated Atmospheres* framework.

Preprocessing

Only relative band power samples of channels 1 and 4 (see Figure 3.12) were considered and dropped if the signal quality indicator (one integer value per channel, provided by the headband) was not sufficient. Histograms were created for each frequency band and each channel, respectively. They were updated with each valid sample being added using the P^2*Algorithm* [Jai85], which allows a dynamic calculation of percentiles and histograms without having to store all observations. Samples falling between the 10th and the 90th percentiles of the histogram were normalized between 0 and 1, whereas other values were considered as outliers and therefore rejected. The subsequent score computations were performed on the mean value of the last 10 samples (further denoted as *processed samples*), hence corresponding to a score update rate of 1 Hz.

Naive score computation

The naive approach for the computation of Focus and Relax scores was purely based on the relative alpha and gamma band powers of the recorded EEG data. Therefore, the *Naive Relax* score was derived from the alpha band by averaging the *processed samples* for channel 1 and 4. Similarly, the *Naive Focus* score corresponded to the *processed samples* in the gamma band, averaged over the two channels. Finally, a 20-point-moving-average low-pass filter was applied to the output values in order to filter out short-time mental state fluctuations.

Entropy-based score computation

In general, entropy serves as a measure for randomness or uncertainty of an information source [Sha48] and is very effective in detecting nonstationary events like peaks and bursts [Ton03]. Hence, in this probabilistic concept the EEG signal was considered as the result of a random process. The *processed samples* for each channel were interpreted as random variables $p_i, i \in \{\delta, \theta, \alpha, \beta, \gamma\}$ emitted by an information source and satisfying the conditions $p_i \geq 0$ and $\sum_i p_i = 1, \forall i$.

Normally, an EEG signal with relatively equal band power distribution has a high degree of randomness and thus exhibits a high entropy. In comparison, an EEG signal with high relative band power in one specific frequency band indicates a decrease of randomness in the signal and therefore results in a lower entropy [Kar10]. Leveraging this, the *Relax* score computation was performed using *processed samples* from alpha and theta frequency bands, whereas gamma and beta band samples were used for the *Focus* score, respectively.

The *Shannon entropy* H_{Sh} [Sha48] can be computed by:

$$H_{Sh} = -\sum_i p_i \cdot log_2(p_i). \tag{3.1}$$

Furthermore, other entropy-related measures were tested for the quantification of being focused or relaxed. A generalization of the Shannon entropy is the *Rényi entropy* H_{Re} [Ren61] of order α, where $\alpha \geq 0$ and $\alpha \neq 1$. It is defined as:

$$H_{Re} = \frac{1}{1-\alpha} \cdot log_2\left(\sum_i p_i^{\alpha}\right). \tag{3.2}$$

Tsallis entropy H_{Ts} [Tsa88] is a non-logarithmic parameterized entropy measure defined as

$$H_{Ts} = \frac{1}{\alpha-1} \cdot \sum_i \left(p_i - p_i^{\alpha}\right). \tag{3.3}$$

In this work, H_{Re} and H_{Ts} entropies of order $\alpha = 3$ were used because they have shown to work well on EEG signals with short-range rhythms [Kar10].

The *Kullback-Leibler divergence* D_{KL} [Kul51] serves as a measure of the difference between two probability density functions p and q and is defined by

$$D_{KL}(p||q) = \sum_i p(i) \cdot log\frac{p(i)}{q(i)}, \tag{3.4}$$

where p and q refer to the *processed samples* of the alpha and theta bands for the Relax score and to the *processed samples* of the gamma and beta bands for the Focus score, respectively.

After computing the entropy measures, they were normalized between 0 and 1 (denoted as H_{norm}). As an increase in being focused or relaxed yields to an decrease of entropy and vice versa, the output values were computed by subtracting H_{norm} from 1. In the same way as for the *Naive* score computation, the entropy-based score values were finally obtained by applying a 20-point-moving-average filter to the output values.

Chapter 4

Evaluation

The goal of this work was to explore the application of smart agents and brain computer interfaces for the interaction with a smart office prototype. Therefore, two experiments were performed individually. In the first experiment (see Section 4.1), different smart agents were analyzed in order to assess whether they are superior to conventional methods, like a smartphone-based graphical user interface (GUI), or a non-conversational voice user interface (VUI). In the second experiment (see Section 4.2), different EEG algorithms were evaluated with regard to their suitability for providing a real-time mental state recognition. Information obtained from the occupants' EEG are used as additional input for the *Sensor Collection Server* for a better estimation of the current *Focus* and *Relax* scores.

4.1 Agents

4.1.1 Experiment Design

The population of interest for *Mediated Atmospheres* is office workers. Therefore, university students, researchers, and local office workers were recruited as subjects for the experiment. The panel consists of N = 33 people (61 % Female), with an average age of 27.5 ± 3.5 (M ± SD). 52 % of the participants named English as their native language. The experiment consists of two parts: *Agent Exploration* and *Task Fulfillment*.

Agent Exploration

For this part of the experiment, participants interacted with all agents (Smartphone, Basic Voice, Advanced Voice, Advanced Text) in randomized order. For every agent, subjects were asked to fulfill the following list of tasks using all possible agent features, with the task order being randomized within the agents:

1. *Turn Mediated Atmospheres off because you're about to get lunch*, followed by *Turn Mediated Atmospheres on after returning from lunch*
2. *Find a scene that has warm color temperature*
3. *Find an indoors scene*
4. *Find a scene that helps you focus*
5. *Find a scene that matches the current weather*

Primarily, the *Agent Exploration* part should make the subjects familiar with the usage of each agent and allow them to explore the range of features. Furthermore, the users' subjective responses were collected in a survey to measure overall usability, perception of intelligence and engagement, as well as perception of trust and control. These measures are explained in detail in Section 4.1.3.

Task Fulfillment

After the first part, the participants should have the same knowledge on how to communicate with every agent, which allows to measure quantitative information about the agent interaction. Each subject was assigned to one randomly selected agent, with nine subjects being assigned to the *Advanced Voice Agent*, eight subjects being assigned to each of the other agents. For the experiment, they were asked to fulfill the following list of tasks utilizing the agent's features in a randomized order:

1. *Find a city scene*
2. *Find a scene that shows mountains*
3. *Find a bright and blue scene*
4. *Find a scene that is warm and shows a forest*
5. *Find a scene that matches the current time of day*
6. *Find a scene that helps you relax after a rough day*

The tasks were selected to have an equal number of tasks for the different features listed in Table 3.2, with (1) and (2) corresponding to *Switch scene [keywords]*, (3) and (4) corresponding to *Switch scene [properties]*, and (5) and (6) corresponding to *Recommend scene [time] / [mental state]*, respectively. For each task, objective measures like fulfillment time, number of interactions, and recognition rate were collected. These measures are explained in detail in Section 4.1.4.

4.1.2 Procedure

Both parts of the experiment were performed in the *Mediated Atmospheres* office space and consecutively in one sitting. Each sitting began with a tutorial, during which the study personnel first described the concept of *Mediated Atmospheres* and its capability of transforming the appearance of the office space using different scenes. Furthermore, the study protocol was explained, followed by an introduction into the different agents, their range of features, and the way of interacting with them. Subsequently, the participants were left alone in the workspace, and a website showing the task list for the currently selected agent guided them through the first part of the experiment. After completing all tasks with one agent, subjects were asked to answer give their feedback in form of a questionnaire and advance with the next agent. Both the order of agents, as well as the order of tasks for every agent were randomized to avoid possible adaption effects. After exploring all agents and giving feedback, the participants were asked to rank the agents according to their preferences.

For the *Task Fulfillment* part, the subjects were left alone in the workspace, again with a website showing a list of tasks to fulfill with the agent assigned to them. The participants were asked to press a button once they successfully accomplished a task to advance with the next one in order to precisely collect task fulfillment times.

4.1.3 Subjective Measures

Overall Usability

For a subjective measure of the usability for each agent, the System Usability Scale (SUS) [Bro96] was used. It consists of a 10 item questionnaire with a five-level Likert scale (1 = Strongly Disagree, 5 = Strongly Agree), and yields a high-level subjective view of the usability as a linear scale of 0 – 100. Based on research, a SUS score of above 68 would be considered *above average*, whereas a SUS score of below 68 would be considered as *below average* [Bro96]. At the end of Part I, participants were asked to directly compare the agents by ranking them according to their preference, with 1 being the most favorite, and 4 being the least favorite agent.

Intelligence and Engagement

The perceived levels of intelligence and engagement were recorded using the questions listed in Section B.1. Additionally, the reasons why the agent were found to be (not) intelligent or engaging, respectively, were of interest.

Trust and Control

As some of the agents recommend scenes based on a certain context, and therefore autonomously make decisions, the perception of trust in the system, and the feeling of being in control, can be affected by the rate of automation error or misautomation. Therefore, the perception of trust and control towards the different agents was measured with a survey proposed by Hossain et al. [Hos13], and adapted to fit the setup of this work. The survey is listed in Section B.2, consisting of an equal number of questions having a negative and positive implication, respectively.

4.1.4 Objective Measures

Fulfillment Time

The fulfillment time for each task was computed as the absolute time for the participants to successfully accomplish one task, i.e. finding a scene that matches the given task in their opinion.

Number of Interactions

For the voice and text agents, the number of interactions was computed by counting how many times the user invoked one of the agents' actions in Table 3.2. For the *Smartphone Application*, the number of interactions was computed by counting how many times the user pressed the *Search for scenes* button per task.

Figure 4.1: Subject wearing a Muse headband during the study.

Recognition Rate

For the voice and text agents, the recognition rate per agent and task was defined as the number of actions correctly mapped to the specific actions, relative to the total number of interactions. For the *Smartphone Application*, the recognition rate was assumed to be 100 % as the user directly changes the scene parameters without any further level of abstraction.

4.2 Brain Computer Interface

4.2.1 Experiment Design

The panel for the evaluation of the BCI consists of $N = 11$ people, with an average age of 28.1 ± 4.6 (M ± SD). The study was conducted in the *Mediated Atmospheres* office space. During the procedure, the subjects wore the Muse headband as proposed by the manufacturer (as shown in Figure 4.1). The data were transmitted to a computer and processed by the *Score Processor*. The study procedure is listed in Figure 4.2 and consisted of multiple tasks. Each of them were associated with one of three possible categories, corresponding to different mental states: *Neutral*, *Focus*, *Relax*. The phases labeled as *Neutral* were used as reference measurements with no specific instructions given, except not to close their eyes. During the phases labeled as *Focus*, the subjects were asked to perform different tasks that were all supposed to generate high levels of mental

Figure 4.2: Study protocol for EEG data acquisition. Each task has a duration of 3 minutes; Red tasks are assigned to the *Focus* category, Green tasks to the *Relax* category.

processing and binding different senses [Cha14, Fer95]. During the phases labeled as *Relax*, the subjects were asked to relax themselves. Additionally, they were able to choose between different nature scenes (e.g. beach, forest, mountain) which had been proven to have a positive effect on achieving a relaxed mental state [Ulr81]. The selected scenes were then displayed to the subjects using a laptop computer.

4.2.2 Measures

The acquired data for each measure were evaluated in order to find the one with best performance. Every data sample was labeled with the associated task during which it was recorded (*Neutral*, *Focus*, *Relax*).

The classification was performed by estimating the probability distributions for each score and measure. Therefore, histograms were computed iteratively for each measure and score, respectively, and were updated in real-time with every new sample using the P^2*-Algorithm*. Subsequently, a binary classification (*Focused / not Focused* for the Focus score and *Relaxed / not Relaxed* for the Relax score, respectively) was applied using a quantile-based threshold. As this method only relied on the data of the subject itself and hence did not require a previously trained classification model, a separation into training and test sets was not necessary and thus, no cross-validation had to be performed.

For determining the best performing measure and the optimal quantile, a Receiver Operating Characteristic (ROC) curve was generated for every measure by computing sensitivity and specificity for quantiles in the interval $[0.05, 0.95]$ with a step size of 0.05. The optimal quantiles for each measure were obtained by selecting the quantile value on the ROC curve with the smallest Euclidean distance to the optimal classifier (1.0 *True Positive Rate* and 0.0 *False Positive Rate*). The best measure for each score was determined by selecting the score measure corresponding to the ROC curve with the highest Area Under the Curve (AUC) value.

Chapter 5

Results

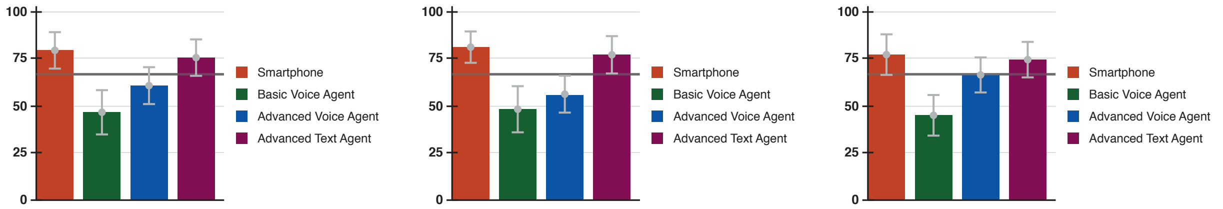
5.1 Agents

5.1.1 Subjective Measures

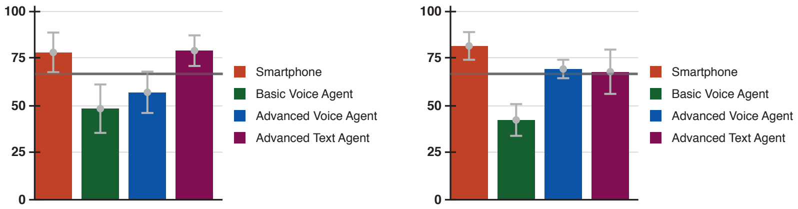
Overall Usability

As visible from Figure 5.1, the *Smartphone Application* achieved on average the highest System Usability Scale (SUS) score (79.8 ± 9.9 (*M* ± *SD*)), followed by the *Advanced Text Agent* (76.3 ± 10.0), the *Advanced Voice Agent* (61.3 ± 10.2), and the *Basic Voice Agent* (47.0 ± 12.0). Only the *Smartphone Application* and the *Advanced Text Agent* have achieved SUS scores above 68, which is considered as being *above average*. Furthermore, the *Advanced Voice Agent* has achieved a significantly higher SUS score for non-native speaker (66.9 ± 9.6) than for native speaker (56.0 ± 10.0). For subjects more than 30 years old, both advanced agents yielded almost the same SUS score, with the *Advanced Voice Agent* achieving a slightly higher SUS score than the *Advanced Text Agent* (70.0 ± 5.0 vs 68.3 ± 12.1). For subjects less than 30 years old, the *Advanced Text Agent* even achieved a higher SUS score than the *Smartphone Application* (79.8 ± 8.3 vs 78.8 ± 10.7).

Results of the participants' ranking of the agents are visualized in Figure 5.2. It shows that, on average, the *Smartphone Application* has achieved the best ranking (1.88 ± 0.42 (*M* ± *SD*)), followed by the *Advanced Text Agent* (2.21 ± 0.51), the *Advanced Voice Agent* (2.61 ± 0.49) and the *Basic Voice Agent* (3.24 ± 0.55). Although the average rating yields the same order among native and non-native speaker, a clear difference can be observed between both groups. Non-native speaker tend to have a clearer aversion towards the *Basic Voice Agent* than native speaker, where the average rankings of the voice and text agents are closer together.

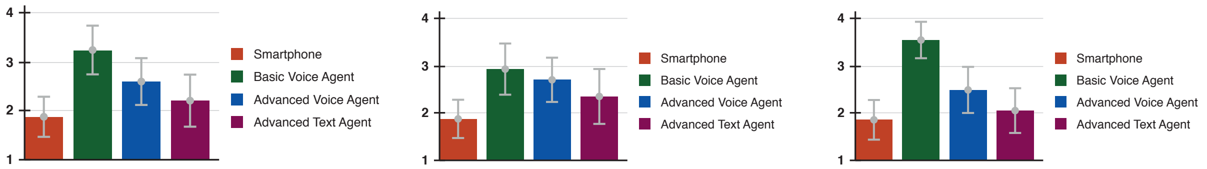


(a) All participants **(b)** Native Speaker **(c)** Non-native Speaker

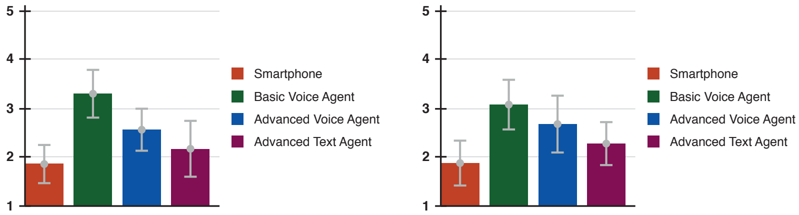


(d) ≤ 30 years old **(e)** > 30 years old

Figure 5.1: Bar charts of System Usability Scale (SUS). The bold, horizontal line represents a SUS score of 68. Error bars have the length of one standard deviation.

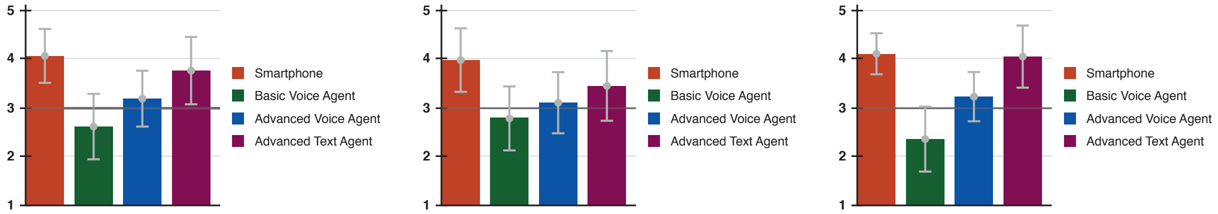


(a) All participants **(b)** Native Speaker **(c)** Non-native Speaker

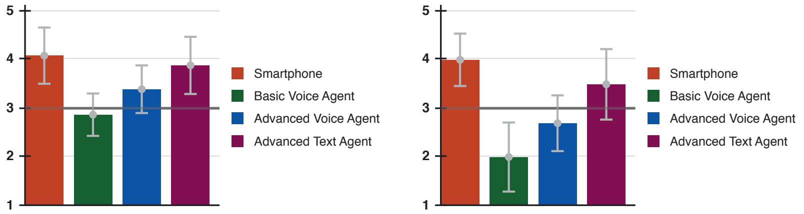


(d) ≤ 30 years old **(e)** > 30 years old

Figure 5.2: Bar charts for the Agent Ranking. Error bars have the length of one standard deviation.

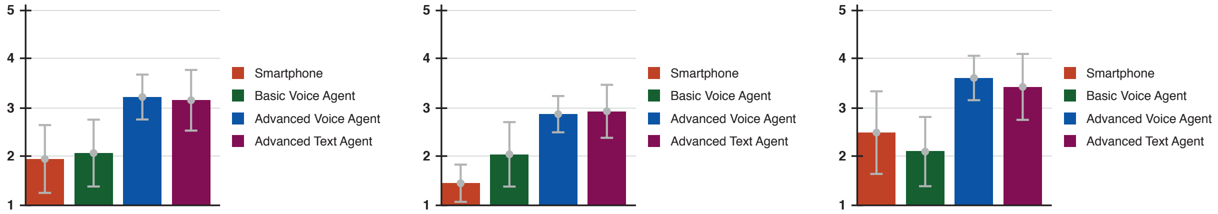


(a) All participants **(b)** Native Speaker **(c)** Non-native Speaker

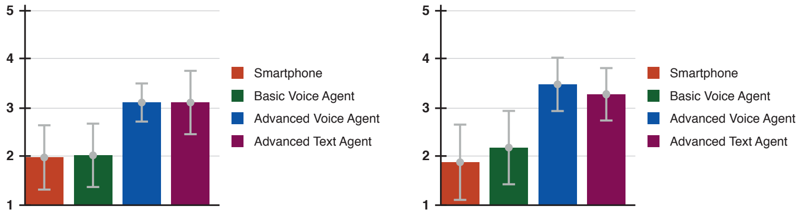


(d) ≤ 30 years old **(e)** > 30 years old

Figure 5.3: Bar charts for User Expectations. A score greater than the bold, horizontal line indicates that the agent exceeded the expectations, a smaller score indicates that the agent did not meet them. Error bars have the length of one standard deviation.

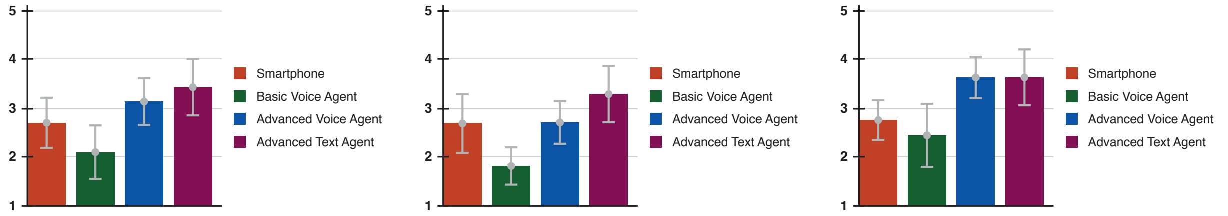


(a) All participants **(b)** Native Speaker **(c)** Non-native Speaker

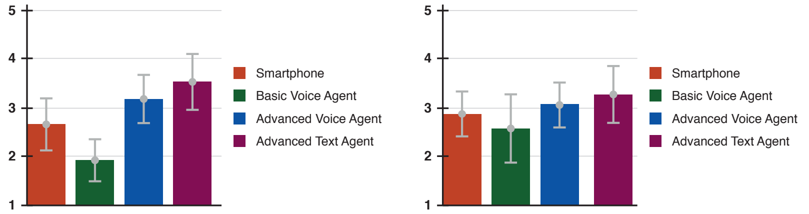


(d) ≤ 30 years old **(e)** > 30 years old

Figure 5.4: Bar charts for Context-Awareness. Scores represent the perceived levels of context-awareness of the agents. Error bars have the length of one standard deviation.



(a) All participants **(b)** Native Speaker **(c)** Non-native Speaker



(d) ≤ 30 years old **(e)** > 30 years old

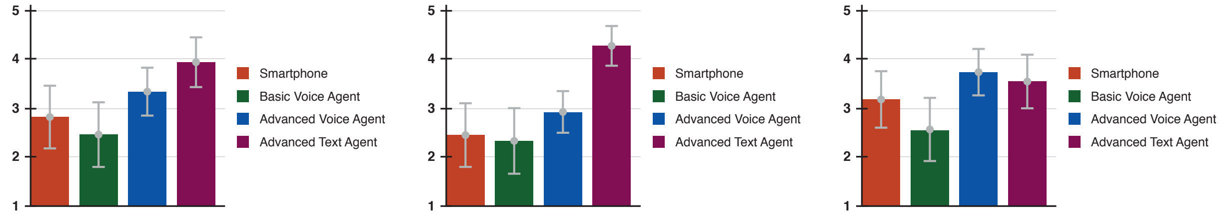
Figure 5.5: Bar charts for Intelligence. Scores represent the perceived levels of intelligence of the agents. Error bars have the length of one standard deviation.

Intelligence and Engagement

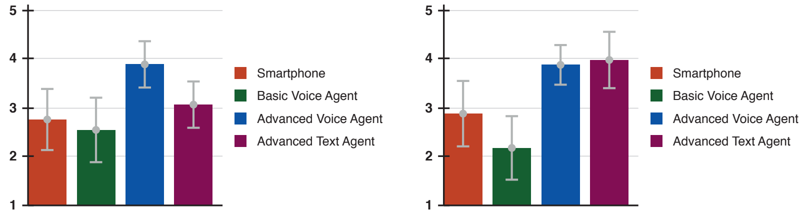
Figure 5.3 shows the response to the question whether the agents met the expectations of the the participants. It suggests that for both native and non-native speaker, the *Basic Voice Agent* is the only condition where the expectations were not met, i.e. where a scoring smaller than 3.0 was achieved. All other systems met the expectations of the participants (*Advanced Voice Agent*) or exceeded them (*Smartphone Application* and *Advanced Text Agent*). For participants more than 30 years old, the *Advanced Voice Agent* also did not meet their expectations, whereas the agent exceeded the expectations of participants less than 30 years old.

The perceived levels of context-awareness are visualized in Figure 5.4, indicating that both advanced agents offered the highest context-awareness. Results also yield that native speaker found the *Basic Voice Agent* to be more context-aware than the *Smartphone*, whereas it was the other way around for non-native speaker.

Figures 5.5 and 5.6 show the perceived levels of intelligence and engagement. Both the *Advanced Text Agent* and the *Advanced Text Agent* achieved the highest levels of intelligence and engagement, with a clear difference between native and non-native speaker.



(a) All participants **(b)** Native Speaker **(c)** Non-native Speaker



(d) $\leq$ 30 years old **(e)** $>$ 30 years old

Figure 5.6: Bar charts for Engagement. Scores represent the perceived levels of engagement of the agents. Error bars have the length of one standard deviation.

Table 5.1: Correlation between positive and negative Trust and Control implications. Pearson's correlation coefficient (PCC) and p-value (p). Significant correlations ($p < 0.05$) are highlighted.

	Smartphone	**Basic Voice**	**Advanced Voice**	Advanced Text
PCC	**-0.711**	**-0.503**	**-0.635**	-0.273
p	$\mathbf{3.54 \cdot 10^{-6}}$	$\mathbf{2.83 \cdot 10^{-3}}$	$\mathbf{7.15 \cdot 10^{-5}}$	0.12

Trust and Control

The results of the perceived levels of trust and control for the agents are listed in Table 5.2. Overall, the *Smartphone Application* achieved the highest levels of trust and control, followed by the *Advanced Text Agent*, the *Advanced Voice Agent* and the *Basic Voice Agent*. Whereas the difference between native and non-native speaker is relatively low for the *Smartphone Application* and the *Basic Voice Agent*, non-native speaker reported a higher level of positive implications the *Advanced Text Agent*. Similarly, the *Advanced Voice Agent* shows higher positive and lower negative implications for trust and control among non-native speakers than for native speakers. Furthermore, Table 5.1 indicates a significant correlation ($p < 0.05$) between positive and negative trust and control implications for all conditions except the *Advanced Text Agent*.

Table 5.2: Perceived levels of Trust and Control. Mean (M) and standard deviation (SD). Significant differences between native and non-native speaker are highlighted.

	Smartphone		Basic Voice		Advanced Voice		Advanced Text	
	M	SD	M	SD	M	SD	M	SD
Positive Implications								
- *All participants*	4.15	0.91	2.78	1.13	3.20	0.85	3.92	0.90
- *Native Speaker*	4.22	0.85	2.74	1.12	**2.91**	0.82	**3.66**	1.03
- *Non-native Speaker*	4.08	0.97	2.83	1.14	**3.50**	0.78	**4.19**	0.62
Negative Implications								
- *All participants*	1.42	0.78	2.34	1.16	2.15	0.87	1.63	0.80
- *Native Speaker*	1.35	0.61	2.33	1.20	**2.31**	0.82	1.66	0.86
- *Non-native Speaker*	1.48	0.93	2.35	1.12	**1.98**	0.89	1.59	0.73

5.1.2 Objective Measures

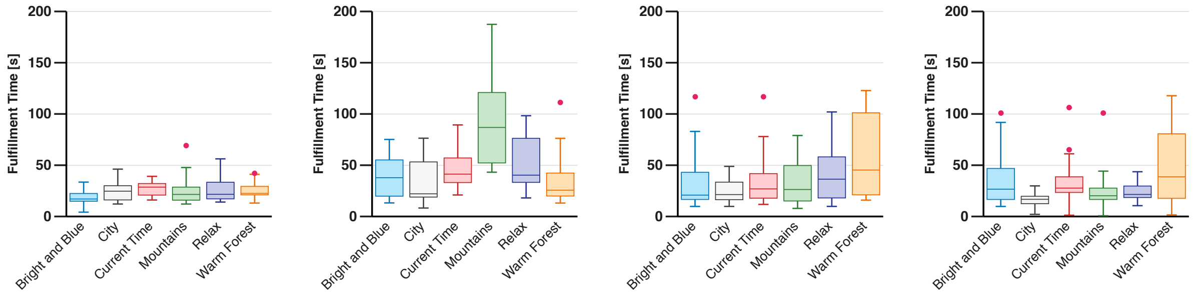
Fulfillment Time

Fulfillment times for the tasks in Part II are visualized as box plots in Figure 5.7. The mean fulfillment times are further listed in Table 5.3. Results yield that the fastest accomplishment of tasks was on average performed with the *Smartphone Application*, and that the fulfillment times show lower variance compared to the other agents. With the *Advanced Text Agent* and the *Advanced Voice Agent*, users needed 1.2 and 1.6 times longer, respectively, whereas the *Basic Voice Agent* had an almost twice as high mean fulfillment time than the *Smartphone Application*.

Furthermore, results show that participants using one of the advanced agents had trouble finding a "warm forest scene" (2.4 and 1.9 times slower, compared to using the smartphone), whereas participants using the *Basic Voice Agent* had particularly difficulties finding a "scene that shows mountains" (3.4 times slower). It is also visible that advanced agents are more susceptible to outliers than other agents. However, on average, three out of the six tasks were performed faster with the *Advanced Text Agent* than with the *Smartphone Application*, and one task was performed faster with the *Advanced Voice Agent*.

Number of Interactions

Figure 5.8 and Table 5.4 show the number of interactions participants needed to fulfill a certain task. In total, the *Advanced Voice* and *Advanced Text* agents required 9 % and 25 % less interactions to fulfill the same tasks than the *Smartphone Application*, whereas participants using the *Basic Voice Agent* required 31 % more interactions. Results also yield



(a) Smartphone **(b)** Basic Voice **(c)** Advanced Voice **(d)** Advanced Text

Figure 5.7: Box plots of Fulfillment Time per task during Part II. Central mark - Median; Box - 25th and 75th percentile; Whiskers - Minimum and Maximum which are not considered outliers; Dot - Outliers.

Table 5.3: Mean Fulfillment Time per task and agent during Part II. Tasks, that were on average fulfilled faster by an agent than with the *Smartphone Application*, are highlighted.

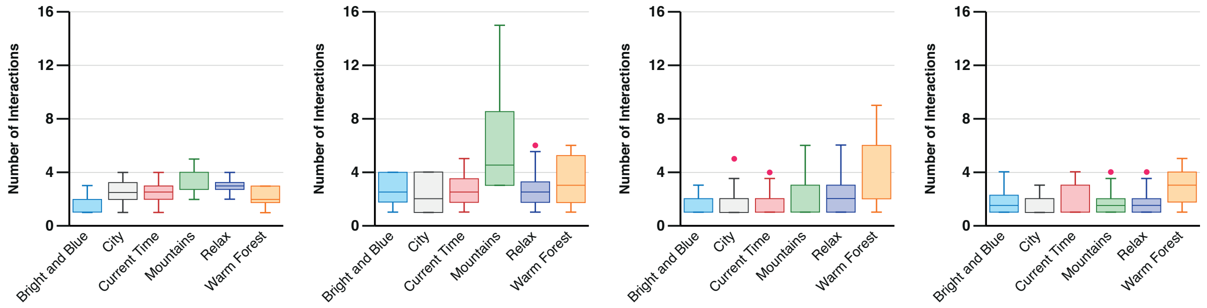
Task	Smartphone	Basic Voice	Advanced Voice	Advanced Text
Bright and Blue	20.5	39.0	38.9	36.3
City	26.6	35.4	**26.0**	**16.6**
Current Time	27.6	47.4	39.4	37.2
Mountains	28.1	96.0	35.0	**27.0**
Relax	26.8	51.8	43.9	**24.8**
Warm Forest	25.5	37.6	59.8	48.4
Total	155.1	307.2	243.0	190.3

that the variance in the necessary interactions is higher for both voice agents compared to the *Advanced Text Agent* and the *Smartphone Application*.

Additionally, the advanced agents required less interactions for five and four out of six tasks, respectively, and the *Basic Voice Agent* still required less interactions for two of the six tasks.

Recognition Rate

The mean recognition rates for each task and agent are listed in Table 5.5. The highest recognition rate was achieved by the *Advanced Text Agent*, followed by the *Advanced Voice Agent* and the *Basic Voice Agent*. Whereas finding a "city scene" achieved high recognition rates throughout all agents, commands leading to finding a "mountain scene" were well recognized by both advanced agents (96.9 % and 88.9 %, respectively), but very poorly



(a) Smartphone **(b)** Basic Voice **(c)** Advanced Voice **(d)** Advanced Text

Figure 5.8: Box plots of Number of Interactions per task during Part II. Central mark - Median; Box - 25th and 75th percentile; Whiskers - Minimum and Maximum which are not considered outliers; Dot - Outliers.

Table 5.4: Number of interactions per task and agent during Part II. Tasks that required on average less interactions using an agent than using the *Smartphone Application* are highlighted.

Task	Smartphone		Basic Voice		Advanced Voice		Advanced Text	
	M	SD	M	SD	M	SD	M	SD
Bright and Blue	1.75	0.66	2.63	1.22	**1.67**	**0.67**	1.88	1.05
City	2.63	0.99	**2.38**	**1.32**	**2.00**	**1.25**	**1.50**	**0.71**
Current Time	2.50	0.87	2.75	1.48	**1.89**	**0.99**	**1.88**	**1.17**
Mountains	3.50	1.00	6.38	4.06	**2.22**	**1.69**	**1.75**	**0.97**
Relax	3.00	0.71	**2.88**	**1.54**	**2.44**	**1.57**	**1.75**	**0.97**
Warm Forest	2.13	0.78	3.38	2.00	3.89	2.77	2.88	1.36
Total	15.50	2.29	20.38	5.98	**14.11**	**5.63**	**11.63**	**2.83**

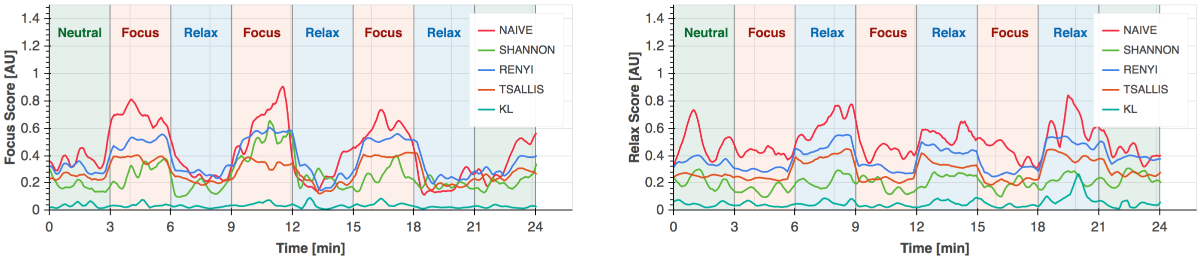
by the *Basic Voice Agent* (55.1 %). In contrast, the *Basic Voice Agent* achieved a higher recognition rate for the task of finding a “warm forest scene” (81.3 %) than the advanced agents (74.6 % and 71.0 %, respectively). Also, the *Advanced Voice Agent* had difficulties recognizing commands to find a “bright and blue scene”, compared to the *Advanced Text Agent* (68.4 % vs. 89.6 %).

5.2 Brain Computer Interface

Focus and Relax scores recorded for one subject while performing the study protocol are visualized in Figure 5.9. The ROC curves for the different measures of Focus and Relax

Table 5.5: Recognition Rate per task and agent during Part II. Minimum and maximum recognition rates for each agent are highlighted.

Task	Smartphone		Basic Voice		Advanced Voice		Advanced Text	
	M	SD	M	SD	M	SD	M	SD
Bright and Blue	100.0	0.0	71.9	23.3	**68.4**	**28.9**	89.6	18.5
City	100.0	0.0	**87.5**	**25.0**	88.1	23.1	93.8	16.5
Current Time	100.0	0.0	64.5	28.7	82.4	26.8	78.1	30.3
Mountains	100.0	0.0	**55.1**	**20.8**	**88.9**	**17.5**	**96.9**	**8.3**
Relax	100.0	0.0	66.6	27.7	74.1	30.5	**96.9**	**8.3**
Warm Forest	100.0	0.0	81.3	24.2	71.0	28.2	**74.6**	**25.2**
Mean	100.0	0.0	71.1	12.7	78.1	14.2	88.5	9.27

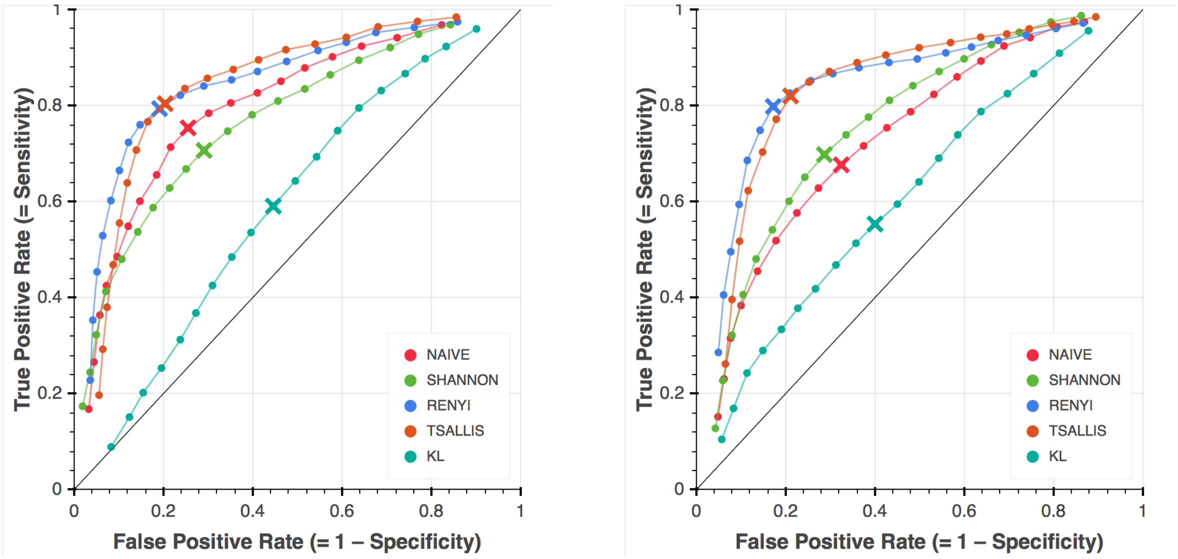


(a) Focus

(b) Relax

Figure 5.9: Course of Focus and Relax score measures. The scores were recorded from a typical subject during the experiment.

scores are shown in Figure 5.10, with the corresponding AUC values in Table 5.6. ROC results indicate that the *Tsallis*-based measure performed best for the Focus score, whereas the *Renyi*-based measure performed best for the Relax score. The optimal quantiles of each measure with their respective sensitivity and specificity are listed in Table 5.7, suggesting that the *Tsallis* score measure offers the highest sensitivity for both *Relax* and Focus scores, whereas the *Renyi* measure achieved the highest specificity. Results also yield that 0.65 is the best performing quantile for the Focus score, compared to 0.55 for the Relax score.



(a) Focus **(b)** Relax

Figure 5.10: ROC curves for Focus and Relax score measures. The best performing quantile for each measure is denoted by an **x**.

Table 5.6: Area Under The Curve (AUC) value for each score measure. The highest AUC values are highlighted.

Measure	Focus score	Relax score
Naive	0.603	0.622
Shannon	0.624	0.614
Renyi	0.702	**0.703**
Tsallis	**0.723**	0.690
KL	0.493	0.493

Table 5.7: Performance evaluation for each measure. The highest sensitivity and specificity values are highlighted.

Measure	**Focus score**	**Relax score**
Naive	*Quantile*: 0.55 *Sensitivity*: 67.7 % *Specificity*: 67.5 %	*Quantile*: 0.50 *Sensitivity*: 75.4 % *Specificity*: 74.4 %
Shannon	*Quantile*: 0.55 *Sensitivity*: 69.8 % *Specificity*: 71.3 %	*Quantile*: 0.50 *Sensitivity*: 70.7 % *Specificity*: 70.9 %
Renyi	*Quantile*: 0.65 *Sensitivity*: 79.8 % *Specificity*: **82.8 %**	*Quantile*: 0.55 *Sensitivity*: 79.3 % *Specificity*: **80.8 %**
Tsallis	*Quantile*: 0.65 *Sensitivity*: **82.0 %** *Specificity*: 78.9 %	*Quantile*: 0.55 *Sensitivity*: **80.4 %** *Specificity*: 79.6 %
KL	*Quantile*: 0.50 *Sensitivity*: 55.3 % *Specificity*: 60.0 %	*Quantile*: 0.50 *Sensitivity*: 59.1 % *Specificity*: 55.4 %

Chapter 6

Discussion

6.1 Agents

6.1.1 Smartphone Application

As apparent from the study results, the smartphone application as the simplest and most familiar system was also the most favorable system. It achieved the highest SUS score, the best overall ranking, and was also the system that exceeded the users' expectations the most. They liked the simple, clear and familiar graphical user interface that facilitated quick and efficient interactions, and especially appreciated the visual preview of scenes. The system provided a comprehensible mapping of user input to scene output with no hidden system actions. Therefore, it also achieved the highest perception of trust and control. On the other hand, the lack of autonomy required users to resolve the tasks and map their requests to scene properties all by themselves. Accordingly, users reported that using the smartphone application felt more like just controlling the output modalities of *Mediated Atmospheres* rather than having an actual interaction and hence made the office space appear inanimate and lonely.

On average, the smartphone application had the smallest task fulfillment time and required the least number of interactions. It outperforms other systems for tasks that could directly be mapped to the sliders of the user interface, such as finding a bright and blue or a warm forest scene. Whereas both advanced agents also offer the feature of filtering scenes based on scene properties, they had trouble correctly mapping a combination of parameters to the appropriate action.

However, users criticized that adjusting the sliders can be time consuming and not as

Table 6.1: Strengths and Weaknesses for each agent.

	Strengths	**Weaknesses**
Smartphone	- Quick, familiar interaction - Predictable actions - Visual scene preview - Efficient filtering of scenes	- No hands-free interaction - No interaction with room - No autonomy - Disrupts workflow easy
Basic Voice	- Hands-free interaction - Manageable feature range, no overtaxing of user	- Rigid, inanimate, time-consuming interaction - No intelligence or context-awareness - Strong familiarity with system required
Advanced Voice	- Hands-free interaction - Engaging, conversational - Abstract scene description possible	- Bad smalltalk handling - Wide gulf of user expectation and experience - False recognition increases with complexity of request
Advanced Text	- Fast interaction, especially in office scenario - Engaging, conversational - Abstract scene description possible - Narrow gulf of user expectation and experience	- No hands-free interaction - Bad smalltalk handling - False recognition increases with complexity of request

intuitive for more abstract tasks, like finding a relaxing scene. Additionally, in an office scenario, users tend to be more distracted, because it disrupts their workflow as they have to switch from the computer to the smartphone and back.

6.1.2 Basic Voice Agent

In general, the basic agent creating the impression of being very rigid and time consuming, and often reminded users of talking to a customer service hotline. Furthermore, the additional problems with voice recognition, causing the user to repeat the same commands over and over, made them feel uncomfortable when interacting with the agent. Combined with the missing intelligence and context-awareness, it led to the lowest overall usability, and also the lowest perceived levels of intelligence and engagement.

However, some users appreciated the limited range of features. This made the basic agent appear structured and led to predictable results – as long as the voice recognition worked correctly. The interaction space was basically reduced to asking the agent to list scenes, and switching to one scene either by its descriptive name or by using the "next scene" command. Therefore, users were able to accomplish some tasks equally fast or even faster with the basic agent than with the advanced agents. For example, for the "Warm Forest" task, they selected a scene based on its descriptive name (e.g. "Forest"), and then switched to another scene using the "next" command until they found a forest scene with warm color temperature.

6.1.3 Advanced Voice Agent

Users reported that the advanced voice agent was very friendly and cooperative in finding a matching scene. The possibility of transforming the workspace using a more open command list, such as providing a high level description of scenes, or by letting the agent autonomously recommend scenes, was appreciated throughout all users, and was often found to be time saving although the net fulfillment time was higher than with the smartphone application.

In addition, the agent tried to conduct a normal conversation by referring to the users by their names, using a more informal language, and providing different replies depending on the current context. During the experiment, some users referred to the agent as "she" instead of "it", because it appeared emotional and engaging, and even made the users laugh occasionally.

However, the SUS score yields that the advanced voice agent was rated worse than the average score, which was mainly due to the voice recognition. It was reported as the main bottleneck that impairs the user experience, creating a gulf between the user expectations towards a seemingly intelligent system, and the inability to correctly recognize commands such as "find a bright and blue scene", which was often falsely recognized as "find a bride and blue scene". Especially native speaker appear to be more critical towards voice recognition. As results indicate throughout the experiment, their judgement is more determined by correct voice recognition and a smooth, natural conversation than by the actual range of features. Therefore, native speakers rated the advanced agent more similar to the basic agent, mentioning that the lack of voice recognition accuracy made it feel dumb. In contrast, non-native speaker noticed a clear difference between both voice agents. Moreover, users reported that the interaction over voice could be time consuming because sometimes it took too long to get a response from the agent, especially when it did not understand the command.

6.1.4 Advanced Text Agent

Compared to the advanced voice agent, the interaction with the advanced text agent was perceived to be more smooth and straightforward as the agent instantly provides a response over text. Similarly, it was on average rated as being more engaging, especially among native speaker. Users reported that the use of emojis was an effective compensation for the missing voice interface. Therefore, the text agent achieved the second highest overall usability and was the only agent with a SUS score greater than the average score.

Both advanced agents required on average less interactions than the smartphone application and can more easily be integrated into a smart office environment. This allows users to interact with the agent while performing other tasks in the meantime. Moreover, the radius of interaction at the workplace is usually more restricted to the computer on the desk, compared to a smart home scenario. Therefore, it might be easier to interact with a text agent over the computer by simply switching to the agent's conversation window rather than having to speak out loud to a voice agent.

During the experiment, it could be observed that users utilized different wordings between the two advanced agents. For instance, when trying to find a relaxing scene, they asked the voice agent "Can you find me a scene that makes me relax?", whereas they used more informal and shorter commands when communicating with the text agent, like "Relaxing scene" or even just "Relax". Users reported that it felt strange *talking* to an agent in such an informal manner, but it did not when they *texted* the system, since it is also usual for them to text their friends in the same way. However, participants emphasized that they were confused by both advanced agents of being conversational on the one side, but offering only limited understanding for small talk and filler language. While the experiment was conducted, API.AI released a new feature that allows to integrate several small talk capabilities into the agent [API17a], which should be added to all agents in future work in order to improve the user experience.

6.2 Brain Computer Interface

For the entropy-based measures, input from two adjacent EEG frequency bands (α and θ for the Relax score, γ and β for the Focus score) were used, whereas the *Naive* measure only used the relative band power from the α and γ bands, respectively. Because the association of certain tasks or mental states with EEG frequency bands are only considered as guidelines and thus might not completely coincide for every person [Cla98], using information from two adjacent frequency bands might increase the classification performance. Additionally, entropy-based approaches for EEG analysis had proven to achieve better results than solely using the relative frequency band power or ratios of relative band powers [Tha13]. Therefore, the results indicate that entropy-based measures *Tsallis* and *Renyi* outperform the *Naive* approach for both Focus and Relax scores. Accordingly, previous work had shown that *Tsallis* and *Renyi* entropies are more effective in characterizing the complexity of the brain [Ton03] than *Shannon* entropy, because the *entropic index* α for H_{Re} and H_{Ts} in Eq. (3.2) and (3.3) can be tuned to emphasize either background activity or bursts in EEG frequency bands. For this application, the entropic index was selected with respect to short-range interactions [Kar10].

The Neutral phase at the beginning of every recording was necessary in order to provide a first estimation of the histograms for the different measures. This allowed the classification of the respective mental state to be performed without using any training data acquired beforehand. The results of the experiment built on the assumption of equal prior probabilities for both Focus and Relax, as the study consisted of equal numbers of tasks associated with each class, leading to a bimodal histogram. When applying this method to other scenarios exhibiting a non-equal class distribution, the histogram would be skewed. Therefore, the experiment design should be adapted in order to obtain a better histogram estimation for mental state recognition. Because subjects were confronted with concrete and demanding Focus tasks, they reported that they felt more focused than they felt relaxed during the Relax phases, where a video of a nature scene was presented to them and they were instructed to relax. Therefore, the AUC values in Table 5.6 indicate that the proposed measures perform better for the Focus score than for the Relax score. Furthermore, as visible in Figure 5.9, the measures achieved higher amplitudes and more distinct peaks during Focus tasks then they did for Relax tasks, also leading to higher quantiles for the Focus score (as denoted in Table 5.7). In general, both *Tsallis* and *Renyi* show appropriate results for a real-time mental state recognition and prove to be a valuable additional input for the *Mediated Atmospheres* framework.

Chapter 7

Conclusion

In this work, different concepts for the interaction with a smart office prototype were introduced, such as a smartphone application, different versions of smart agents, and a brain computer interface. When using an agent, users can dynamically change the appearance of their workspace either by switching to particular scenes, by describing scenes in an abstract way, or by having the agent recommend scenes based on a certain context. With the brain computer interface, users can project their mental state into their environment by mapping the current levels of Focus and Relax to an appropriate scene.

For the smart agents, an experiment was conducted with the aim to evaluate their usability and to analyze whether the agents created different perceptions on the user. Results show that the *Smartphone Application* as the most familiar system is also the system that achieved the highest usability and the best overall ranking. It is followed by the *Advanced Text Agent* and the *Advanced Voice Agent*, which were both perceived as very friendly and engaging conversational agents that facilitate finding the right scenes. However, the results also indicated that especially the *Advanced Voice Agent* created a clearly different perception among native and non-native speaker. It could be observed that non-native speaker more considered the range of features into their decision, whereas the overall impression of native speaker was more influenced by correct voice recognition and a smooth, natural conversation. In general, the findings show that all agents have their strengths and weaknesses, with the *Advanced Text Agent* offering a well appreciated trade-off between a quick and easy interaction, a natural, engaging and entertaining conversation, and an open command list with the possibility to let the agents recommend scenes based on an abstract description.

Additionally, the application of a brain computer interface based on a wearable EEG headband was evaluated in a separate experiment. Therefore, algorithms for providing real-time mental state recognition and a classification of being focused or relaxed were presented. A data-driven approach was used by estimating the probability distribution for each person and updating it in real-time. The classification was then performed by using quantiles of the estimated histogram as threshold. A *Naive* approach was compared to different entropy-based approaches, showing that the measures using *Renyi* and *Tsallis* entropy performed better than the *Naive* or the *Shannon* measure. By finding a trade-off between sensitivity and specificity, the best quantile was 0.65 for the Focus score with a sensitivity of 82.0 % (*Tsallis*) and a specificity of 82.8 % (*Renyi*). The best performing quantile for the Relax score was 0.55, with a sensitivity of 80.4 % (*Tsallis*) and 80.8 % (*Renyi*). Therefore, results show that a mental state recognition can be performed in real-time and without building a classification model using previously recorded training data. It also indicates that alternative entropy measures other than Shannon entropy or KL divergence have to be established in order to assess specific applications. In general, it proves the possibility of applying this system as a brain computer interface for a context-aware system, such as *Mediated Atmospheres*.

Appendix A

Patents

A.1 Method for controlling device by using brain waves

Publication Number	US12716425
Date of Publication	Mar. 3, 2010
Inventor(s)	**Yoshihisa Terada**, Osaka (JP); **Koji Morikawa**, Kyoto (JP)
Assignee	**Panasonic Corp**, Osaka (JP)
Abstract	The control method for a device includes steps of: presenting a visual stimulation concerning a manipulation menu for a device; measuring event-related potentials after the visual stimulation is presented, where event-related potentials based on a timing of presenting the visual stimulation as a starting point are measured from a potential difference between each of electrodes and at least one reference electrode respectively worn on a face and in an ear periphery of a user; from each of the measured event-related potentials, extracting electroencephalogram data which is at 5 Hz or less and contains a predetermined time section, and combining the extracted electroencephalogram data into electroencephalogram characteristic data.

A.2 Electroencephalogram interface system

Publication Number	US12955016
Date of Publication	Mar. 24, 2011
Inventor(s)	**Yoshihisa Terada**, Tokyo (JP); **Koji Morikawa**, Tokyo (JP)
Assignee	**Panasonic Corp**, Osaka (JP)
Abstract	An eyeglass-type electroencephalogram interface system is worn on the head of a user. The system includes: an output section for presenting a visual stimulation to the user; an ear electrode portion disposed at a position coming in contact with an ear of the user when the system is worn; a facial electrode portion disposed at a position coming in contact with the face below a straight line connecting an external canthus and an internal canthus of an eye of the user, such that the mass of the system is supported at the position, when the system is worn; and an electroencephalogram measurement and determination section for measuring an event-related potential on the basis of a potential difference between the ear electrode portion and the facial electrode portion based on the visual stimulation being presented by the output section as a starting point.

A.3 Wearable computing apparatus and method

Publication Number	US14216925
Date of Publication	Mar. 17, 2014
Inventor(s)	**Christopher Allen Aimone**, Toronto (CA) **Ariel Stephanie Garten**, Toronto (CA); **Trevor Coleman**, Toronto (CA)
Assignee	**InteraXon Inc.**, Toronto (CA)
Abstract	A method is provided, performed by a wearable computing device comprising at least one bio-signal measuring sensor, the at least one bio-signal measuring sensor including at least one brainwave sensor, comprising: acquiring at least one bio-signal measurement from a user using the at least one bio-signal measuring sensor, the at least one bio-signal measurement comprising at least one brainwave state measurement; processing the at least one bio-signal measurement, including at least the at least one brainwave state measurement, in accordance with a profile associated with the user; determining a correspondence between the processed at least one bio-signal measurement and at least one predefined device control action; and in accordance with the correspondence determination, controlling operation of at least one component of the wearable computing device, such as modifying content displayed on a display of the wearable computing device. Various types of bio-signals, including brainwaves, may be measured and used to control the device in various ways.

A.4 Conversational interface agent

Publication Number	US7019749
Date of Publication	Jul. 31, 2003
Inventor(s)	**Baining Guo**, Bellevue, WA (US); **Bo Zhang**, Beijing (CN); **Heung-Yeung Shum**, Beijing (CN)
Assignee	**Microsoft Corp**, Redmond, WA (US)
Abstract	A video rewrite technique for rendering a talking head or agent completely simulates a conversation by including a waiting or listening state. Smooth transitions are provided to and from a talking state.

A.5 System and method for a cooperative conversational voice user interface

Publication Number	US8073681
Date of Publication	Apr. 17, 2008
Inventor(s)	**Larry Baldwin**, Maple Valley, WA (US); **Tom Freeman**, Mercer Island, WA (US); **Michael Tjalve**, Bellevue, WA (CN); **Blane Ebersold**, Seattle, WA (US); **Chris Weider**, Everett, WA (US)
Assignee	**VoiceBox Technologies Inc.**, Bellevue, WA (US)
Abstract	A cooperative conversational voice user interface is provided. The cooperative conversational voice user interface may build upon short-term and long-term shared knowledge to generate one or more explicit and/or implicit hypotheses about an intent of a user utterance. The hypotheses may be ranked based on varying degrees of certainty, and an adaptive response may be generated for the user. Responses may be worded based on the degrees of certainty and to frame an appropriate domain for a subsequent utterance. In one implementation, misrecognitions may be tolerated, and conversational course may be corrected based on subsequent utterances and/or responses.

Appendix B

Survey Questions

B.1 Measure of Intelligence and Engagement

1. Did the system meet your expectations regarding its intelligence?
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)
2. Why did the system meet or not meet your expectations?
3. I think the system is intelligent
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)
4. Why do you think the system is or is not intelligent?
5. I think the system is engaging
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)
6. Why do you think the system is or is not engaging?
7. I think the system is aware of the current context
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)
8. Why do you think the system is or is not aware of the current context?

B.2 Measure of Trust and Control

Positive implications

1. The system's actions are deceptive
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

2. I am suspicious of the system's action or output
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

3. I am wary of the system
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

4. The system is intruding
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

Negative implications

5. I am confident in the system's actions
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

6. All actions of the system are comprehensible
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

7. The system is reliable
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

8. I feel in control while using the system
   (Strongly Disagree) -2 | -1 | 0 | +1 | +2 (Strongly Agree)

Appendix C

Supplementary Material

C.1 Android Client for Amazon DynamoDB

```

@DynamoDBTable(tableName="SceneLibrary")
public class Scene {
  private String name, primaryColor;
  private int id, brightness, colorTemp;

  @DynamoDBIndexRangeKey(attributeName="name")
  public void setName(String name) { this.name = name; }

  @DynamoDBIndexRangeKey(attributeName="name")
  public String getName() { return name; }

  @DynamoDBHashKey(attributeName="id")
  public void setId(int id) { this.id = id; }

  @DynamoDBHashKey(attributeName="id")
  public int getId() { return idid }

  @DynamoDBAttribute(attributeName="brightness")
  public void setBrightness(int brightness) {this.brightness = brightness;}

  @DynamoDBAttribute(attributeName="brightness")
  public double getBrightness() { return brightness; }


```

```
  @DynamoDBAttribute(attributeName="color_temperature")
  public void setColorTemp(int colorTemp) { this.colorTemp = colorTemp; }

  @DynamoDBAttribute(attributeName="color_temperature")
  public double getColorTemp() { return colorTemp; }

  @DynamoDBAttribute(attributeName="primary_color")
    public void setPrimaryColor(String primaryColor) {
    this.primaryColor = primaryColor;
  }

  @DynamoDBAttribute(attributeName="primary_color")
    public void getPrimaryColor() { return primaryColor; }

}


private DynamoDBMapper mMapper = new DynamoDBMapper(dynamoDbClient);
public void getSceneList(SceneDbCallback callback) {
  final DynamoDbScanExpression scanExpr = new DynamoDbScanExpression();
  new Thread(new Runnable() {
    public void run() {
      PaginatedScanList<Scene> result = mMapper.scan(Scene.class, scanExpr);
      Scene[] sceneList = result.toArray(new Scene[result.size()]);
      callback.onSceneListLoaded(sceneList);
    }
  }).start();
}
```

Listing C.1: DynamoDB client in Android

C.2 Advanced Agent Dialogs

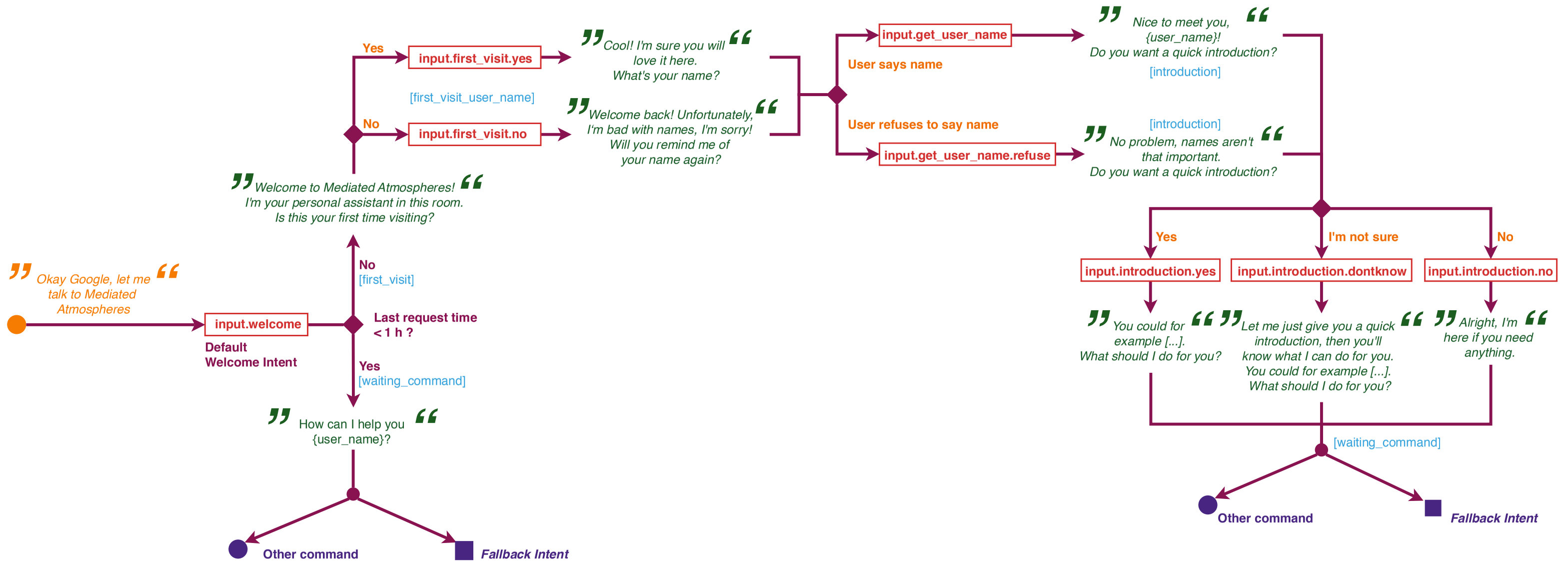


Figure C.1: Flow chart of Welcome Dialog. The dialog is triggered when the user invoked the *Mediated Atmospheres* agent. *Red – Invoked action; Purple – System function; Blue – Current context; Orange – User Input; Green – Agent response.*

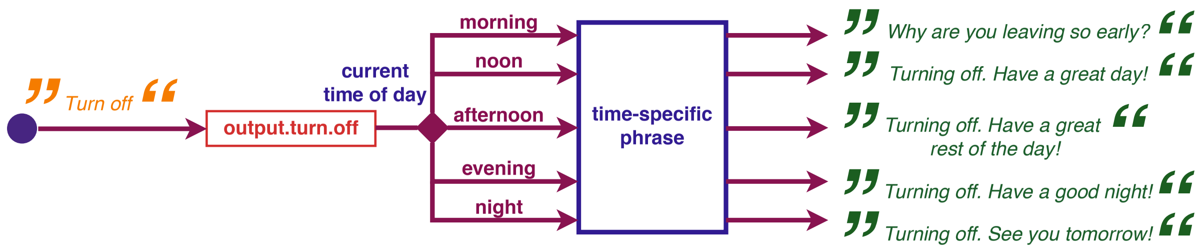


Figure C.2: Flow chart of Turn Off Dialog. The dialog is triggered when the user turns *Mediated Atmospheres* off.

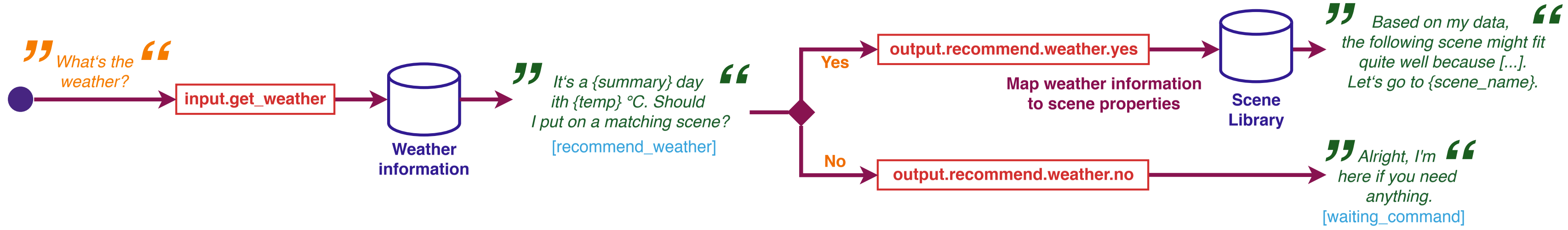


Figure C.3: Flow chart of Weather Dialog. The dialog is triggered when the user asks the agent about the current weather.

List of Figures

List of Tables

Bibliography

[Abo99] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith, and Pete Steggles. Towards a better understanding of context and context-awareness. In *International Symposium on Handheld and Ubiquitous Computing*, pages 304–307. Springer, 1999.

[Abr08] Shadi Abras, Stéphane Ploix, Sylvie Pesty, and Mireille Jacomino. A multi-agent home automation system for power management. *Informatics in Control Automation and Robotics*, pages 59–68, 2008.

[Abu16] Maher Abujelala, Cheryl Abellanoza, Aayush Sharma, and Fillia Makedon. Brain-EE: Brain enjoyment evaluation using commercial EEG headband. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, page 33. ACM, 2016.

[Ala16] Alper T Alan, Enrico Costanza, Sarvapali D Ramchurn, Joel Fischer, Tom Rodden, and Nicholas R Jennings. Tariff agent: interacting with a future smart energy system at home. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 23(4):25, 2016.

[Ald10] Matthew Aldrich, Nan Zhao, and Joseph Paradiso. Energy efficient control of polychromatic solid state lighting using a sensor network. In *SPIE Optical Engineering+ Applications*, pages 778408–778408. International Society for Optics and Photonics, 2010.

[Amo16] Judith Amores, Xavier Benavides, and Pattie Maes. PsychicVR: Increasing mindfulness by using Virtual Reality and Brain Computer Interfaces. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2–2. ACM, 2016.

[And15] Sally Andrews, David A Ellis, Heather Shaw, and Lukasz Piwek. Beyond self-report: Tools to compare estimated and real-world smartphone use. *PloS one*, 10(10):e0139004, 2015.

[API17a] API.AI. API.AI – Documentation – Small Talk. https://docs.api.ai/docs/small-talk-migration, 2017. accessed 2017-05-09.

[API17b] API.AI. API.AI – Integrations. https://docs.api.ai/docs/integrations, 2017. accessed 2017-05-06.

[API17c] API.AI. API.AI – Introduction. https://docs.api.ai/docs/key-concepts, 2017. accessed 2017-04-06.

[Bag03] Faruk Bagci, Jan Petzold, Wolfgang Trumler, and Theo Ungerer. Ubiquitous mobile agent system in a P2P-network. In *UbiSys-Workshop at the Fifth Annual Conference on Ubiquitous Computing, Seattle, USA*, pages 12–15. Citeseer, 2003.

[Bas16] Pouya Bashivan, Irina Rish, and Steve Heisig. Mental State Recognition via Wearable EEG. *arXiv preprint arXiv:1602.00985*, 2016.

[Ben14] Hrvoje Benko, Andrew D Wilson, and Federico Zannier. Dyadic projected spatial augmented reality. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 645–655. ACM, 2014.

[Bre90] Susan Brennan. Conversation as direct manipulation: An iconoclastic view. In Brenda Laurel and S Joy Mountford, editors, *The Art of Human-Computer Interface Design*. Addison-Wesley Longman Publishing Co., Inc., 1990.

[Bro96] John Brooke. SUS – A quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[Cha14] Hung-Cheng Chang, Stephen Grossberg, and Yongqiang Cao. Where's Waldo? How perceptual, cognitive, and emotional brain processes cooperate during learning to categorize and find desired objects in a cluttered scene. *Frontiers in integrative neuroscience*, 8:43, 2014.

[Che13] Nancy H Chen, Jason Nawyn, Maria Thompson, Julie Gibbs, and Kent Larson. Context-aware tunable office lighting application and user response. In *SPIE Optical Engineering+ Applications*, pages 883507–883507. International Society for Optics and Photonics, 2013.

[Chu10] Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. Who is tweeting on Twitter: human, bot, or cyborg? In *Proceedings of the 26th annual computer security applications conference*, pages 21–30. ACM, 2010.

[Cla98] JW Clark. The origin of biopotentials. *Medical Instrumentation, Application and Design, 3rd ed., John Wiley and Sons Inc., New York*, pages 121–182, 1998.

[Cla02] Adam R Clarke, Robert J Barry, Rory McCarthy, and Mark Selikowitz. Eeg analysis of children with attention-deficit/hyperactivity disorder and comorbid reading disabilities. *Journal of Learning Disabilities*, 35(3):276–285, 2002.

[Coo03] Diane J Cook, Michael Youngblood, Edwin O Heierman, Karthik Gopalratnam, Sira Rao, Andrey Litvin, and Farhan Khawaja. MavHome: An agent-based smart home. In *Pervasive Computing and Communications, 2003.(PerCom 2003). Proceedings of the First IEEE International Conference on*, pages 521–524. IEEE, 2003.

[Dan08] Maria Danninger and Rainer Stiefelhagen. A context-aware virtual secretary in a smart office environment. In *Proceedings of the 16th ACM international conference on Multimedia*, pages 529–538. ACM, 2008.

[D'H15] Luis Fernando D'Haro, Seokhwan Kim, Kheng Hui Yeo, Ridong Jiang, Andreea I Niculescu, Rafael E Banchs, and Haizhou Li. CLARA: A multifunctional virtual agent for conference support and touristic information. In *Natural Language Dialog Systems and Intelligent Assistants*, pages 233–239. Springer, 2015.

[DlT15] Fernando De la Torre, Wen-Sheng Chu, Xuehan Xiong, Francisco Vicente, Xiaoyu Ding, and Jeffrey Cohn. Intraface. In *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, volume 1, pages 1–8. IEEE, 2015.

[Duf09] Jeanne F Duffy and Charles A Czeisler. Effect of light on human circadian physiology. *Sleep medicine clinics*, 4(2):165–177, 2009.

[Emo17] Emotiv. Software for Wireless EEG Systems - Emotiv. https://www.emotiv.com/software, 2017. accessed 2017-02-01.

[Fel10] Mark Feldmeier and Joseph A Paradiso. Personalized HVAC control system. In *Internet of Things (IOT), 2010*, pages 1–8. IEEE, 2010.

[Fer95] Thalía Fernández, Thalía Harmony, Mario Rodríguez, Jorge Bernal, Juan Silva, Alfonso Reyes, and Erzsébet Marosi. EEG activation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalography and clinical Neurophysiology*, 94(3):175–182, 1995.

[Fer16] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016.

[Gel12] Tom Geller. Talking to machines. *Communications of the ACM*, 55(4):14–16, 2012.

[Goo16] Google. One assistant. https://assistant.google.com/one-assistant/, 2016. accessed 2017-05-11.

[Hos13] M Anwar Hossain, Ali Asghar Nazari Shirehjini, Abdullah S Alghamdi, and Abdulmotaleb El Saddik. Adaptive interaction support in ambient-aware environments based on quality of context information. *Multimedia tools and applications*, 67(2):409–432, 2013.

[How84] Edgar Howarth and Michael S Hoffman. A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1):15–23, 1984.

[Int17] InteraXon. Muse Research Tools – Available Data. http://developer.choosemuse.com/research-tools/available-data, 2017. accessed 2017-01-10.

[Ish98] Hiroshi Ishii, Craig Wisneski, Scott Brave, Andrew Dahley, Matt Gorbet, Brygg Ullmer, and Paul Yarin. ambientROOM: integrating ambient media with architectural space. In *CHI 98 conference summary on Human factors in computing systems*, pages 173–174. ACM, 1998.

[Jai85] Raj Jain and Imrich Chlamtac. The P^2 Algorithm for Dynamic Calculation of Quantiles and Histograms without storing Observations. *Communications of the ACM*, 28(10):1076–1085, 1985.

[Jap09] Budi Thomas Jap, Sara Lal, Peter Fischer, and Evangelos Bekiaris. Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*, 36(2):2352–2359, 2009.

[Kai05] Anssi Kainulainen, Markku Turunen, Jaakko Hakulinen, Esa-pekka Salonen, and Perttu Prusi. A speech-based and auditory ubiquitous office environment. In *Proc. of 10th International Conference on Speech and Computer (SPECOM 2005*. Citeseer, 2005.

[Kal16] Raghuram Kalyanam and Sabine Hoffmann. Visual and thermal comfort with electrochromic glass using an adaptive control strategy. In *Proceedings of the 15th International Radiance Workshop (IRW)*, 2016.

[Kar10] Sibsambhu Kar, Mayank Bhagat, and Aurobinda Routray. EEG signal analysis for the assessment and quantification of driver`s fatigue. *Transportation research part F: traffic psychology and behaviour*, 13(5):297–306, 2010.

[Kel05] Simon P Kelly, Edmund C Lalor, Ciarán Finucane, Gary McDarby, and Richard B Reilly. Visual spatial attention control in an independent brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 52(9):1588–1596, 2005.

[Kid99] Cory D Kidd, Robert Orr, Gregory D Abowd, Christopher G Atkeson, Irfan A Essa, Blair MacIntyre, Elizabeth Mynatt, Thad E Starner, and Wendy Newstetter. The aware home: A living laboratory for ubiquitous computing research. In *International Workshop on Cooperative Buildings*, pages 191–198. Springer, 1999.

[Kle99] George H Klem, Hans Otto Lüders, HH Jasper, C Elger, et al. The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3):3–6, 1999.

[Kli99] Wolfgang Klimesch. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2):169–195, 1999.

[Kop05] Stefan Kopp, Lars Gesellensetter, Nicole C Krämer, and Ipke Wachsmuth. A conversational agent as museum guide–design and evaluation of a real-world application. In *International Workshop on Intelligent Virtual Agents*, pages 329–343. Springer, 2005.

[Kul51] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[Lug16] Ewa Luger and Abigail Sellen. Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5286–5297. ACM, 2016.

[Lum13] Sigma Luminous. The Circadian Rhythm and Color Temperature. https://sigmaluminous.com/the-circadian-rhythm-and-color-temperature, 2013. accessed 2017-04-15.

[Mae94] Pattie Maes et al. Agents that reduce work and information overload. *Communications of the ACM*, 37(7):30–40, 1994.

[Mag11] Remco Magielse, Philip Ross, Sunder Rao, Tanir Ozcelebi, Paola Jaramillo, and Oliver Amft. An interdisciplinary approach to designing adaptive lighting environments. In *Intelligent Environments (IE), 2011 7th International Conference on*, pages 17–24. IEEE, 2011.

[Med17] Medtronic. Zephyr bioharness 3. https://www.zephyranywhere.com/system/overview, 2017. accessed 2017-05-10.

[Men14] Sarah Mennicken, Jo Vermeulen, and Elaine M Huang. From today's augmented houses to tomorrow's smart homes: New directions for home automation research. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 105–115. ACM, 2014.

[Men16] Sarah Mennicken, Oliver Zihler, Frida Juldaschewa, Veronika Molnar, David Aggeler, and Elaine May Huang. It's like living with a friendly stranger: Perceptions of personality traits in a smart home. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 120–131. ACM, 2016.

[Moz98] Michael C Mozer. The neural network house: An environment hat adapts to its inhabitants. In *Proc. AAAI Spring Symp. Intelligent Environments*, volume 58, 1998.

[Moz05] Michael Mozer, Robert Dodier, Debra Miller, Marc Anderson, J Anderson, D Bertini, M Bronder, M Colagrosso, R Cruickshank, B Daugherty, et al. The adaptive house. In *IEE Seminar Digests*, volume 11059, pages v1–39. IET, 2005.

[Mül08] Klaus-Robert Müller, Michael Tangermann, Guido Dornhege, Matthias Krauledat, Gabriel Curio, and Benjamin Blankertz. Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring. *Journal of neuroscience methods*, 167(1):82–90, 2008.

[Neu17] NeuroSky. EEG Sensors - EEG Headsets - NeuroSky. http://neurosky.com/biosensors/eeg-sensor/biosensors, 2017. accessed 2017-05-20.

[oG17a] Actions on Google. Actions on Google – app basics. https://developers.google.com/actions/apiai/, 2017. accessed 2017-04-07.

[oG17b] Actions on Google. Design Principles and Methodology – Picking your voice. https://developers.google.com/actions/design/principles, 2017. accessed 2017-04-16.

[Pej16] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. Room2Room: Enabling life-size telepresence in a projected augmented reality environment. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 1716–1725. ACM, 2016.

[Per14] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials*, 16(1):414–454, 2014.

[Pic16] Sundar Pichai. I/O: Building the next evolution of Google. https://blog.google/products/assistant/io-building-next-evolution-of-google, 2016. accessed 2017-04-07.

[Ras98] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 179–188. ACM, 1998.

[Ren61] Alfred Renyi. On measures of entropy and information. In *Fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 547–561, 1961.

[Ros15] Marco Rossi, Ashish Pandharipande, David Caicedo, Luca Schenato, and Angelo Cenedese. Personal lighting control with occupancy and daylight adaptation. *Energy and Buildings*, 105:263–272, 2015.

[Ros16] Seth Rosenberg. How To Build Bots for Messenger. https://developers.facebook.com/blog/post/2016/04/12/bots-for-messenger, 2016. accessed 2017-04-07.

[Row85] N Rowland, MJ Meile, S Nicolaidis, et al. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. 1985.

[Sch77] Daniel L Schacter. EEG theta waves and psychological phenomena: a review and analysis. *Biological psychology*, 5(1):47–82, 1977.

[Sch10] Holger Schnädelbach, Kevin Glover, and Ainojie Alexander Irune. ExoBuilding: Breathing life into architecture. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, pages 442–451. ACM, 2010.

[Sha48] Claude Elwood Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[Sim11] Michael Simon, Eike A Schmidt, Wilhelm E Kincses, Martin Fritzsche, Andreas Bruns, Claus Aufmuth, Martin Bogdan, Wolfgang Rosenstiel, and Michael Schrauf. EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clinical Neurophysiology*, 122(6):1168–1178, 2011.

[Siv12] Swaminathan Sivasubramanian. Amazon dynamoDB: a seamlessly scalable non-relational database service. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, pages 729–730. ACM, 2012.

[Smi15] Aaron Smith. U.S. Smartphone Use in 2015. http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015, 2015. accessed 2017-04-03.

[Str16] Carolin Straßmann, Astrid Rosenthal von der Pütten, Ramin Yaghoubzadeh, Raffael Kaminski, and Nicole Krämer. The Effect of an Intelligent Virtual Agent's Nonverbal Behavior with Regard to Dominance and Cooperativity. In *International Conference on Intelligent Virtual Agents*, pages 15–28. Springer, 2016.

[Sub16] VS Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. The DARPA Twitter bot challenge. *Computer*, 49(6):38–46, 2016.

[Tha13] Nitish V. Thakor and David L. Sherman. *EEG signal processing: Theory and applications*, pages 259–303. Springer US, 1 2013.

[Tom08] Martin Tomitsch, Thomas Grechenig, Andrew Vande Moere, and Sheldon Renan. Information sky: Exploring the visualization of information on architectural ceilings. In *Information Visualisation, 2008. IV'08. 12th International Conference*, pages 100–105. IEEE, 2008.

[Ton03] Shanbao Tong, Anastasios Bezerianos, Amit Malhotra, Yisheng Zhu, and Nitish Thakor. Parameterized entropy analysis of EEG following hypoxic–ischemic brain injury. *Physics Letters A*, 314(5):354–361, 2003.

[Tsa88] Constantino Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, 52(1-2):479–487, 1988.

[Tur50] Alan M Turing. Computing machinery and intelligence. *Mind*, 59(236):433–460, 1950.

[Ulr81] Roger S Ulrich. Natural versus urban scenes – some psychophysiological effects. *Environment and behavior*, 13(5):523–556, 1981.

[Wei66] Joseph Weizenbaum. Eliza – A Computer Program For the Study of Natural Language Communication Between Man And Machine. *Communications of the ACM*, 9(1):36–45, 1966.

[Zha14] Chi Zhang, Hong Wang, and Rongrong Fu. Automated detection of driver fatigue based on entropy and complexity measures. *IEEE Transactions on Intelligent Transportation Systems*, 15(1):168–177, 2014.

[Zha17] Nan Zhao, Asaph Azaria, and Joseph A. Paradiso. Mediated Atmospheres: Understanding the Effects of Multimodal Mediated Environments. In *PACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, pages 0–0. ACM, 2017.