



Avoiding Mistakes in Medical High-Tech Treatments and E-Commerce Applications – a Salutory UX-Research Innovation

Christina Miclau¹(✉), Oliver Gast², Julius Hertel³, Anja Wittmann³,
Achim Hornecker³, and Andrea Mueller²

¹ Hochschule Offenburg, University of Applied Sciences,
Badstrasse 24, 77652 Offenburg, Germany
christina.miclau@hs-offenburg.de

² User Interface Design GmbH, Wilhelm-Bleyle-Straße 10-12,
71636 Ludwigsburg, Germany

³ Dr. Hornecker Software-Entwicklung und IT-Dienstleistungen,
Leo-Wohleb-Strasse 6, 79098 Freiburg im Breisgau, Germany

Abstract. Medical devices accompany our everyday life and come across in situations of worse condition, in significant moments concerning the health or during routine checkups. To ensure flawless operations and error-free results it is essential to test applications and devices. High risks for patient's health come with operating errors [33] so that the presented research project, called Professional UX, identifies signals and irritations caused by the interaction with a certain device by analyzing mimic, voice and eye tracking data during user experience tests. Besides, this paper will provide information on typical errors of interactive applications which are based on an empirical lab-based survey and the evaluated results achieved. The pictured proceeding of user experience tests and the following analysis can also be applied to other fields and serves as a support for the optimization of products and systems.

Keywords: Medical technology · User experience · Operation errors · High risk systems

1 Introduction: UX-Challenges for Medical High-Tech Treatments and E-Commerce

The user experience (UX) of graphical user interfaces is becoming an increasingly important prerequisite for the acceptance of software in the age of digitalization. In many areas, easily and intuitively comprehensible user interfaces are also a qualification for the use of software systems, for example in the field of medical devices or in automotive and application scenarios in the area of occupational safety. Often UX is already part of industry standards in these areas. But also in less critical application areas like e-commerce, UX is becoming more and more an economical factor.

However, at the moment the assessment of user experience is still mainly based on subjective criteria, which are the results of surveys and experience of professionals, despite guidelines such as DIN EN 62366. This project uses artificial intelligence

techniques to develop objective measurements for determining user experience building on non-invasive methods such as eye movements, facial expressions and measurement in the change of voice. From these measurements there will be derived criteria for evaluating user interfaces, such as anger, confusion or stress.

2 Methods to Identify User Irritations for Avoiding Mistakes

Unfortunately, there is still no standardized value for user experience in case of the optimization of interactive applications or devices. The various number of methods makes it difficult to compare results because of the different approaches.

According to Sarodnick and Brau [1] two kinds of procedures can be used for measuring user experience, see Fig. 1 and its description in Table 1.

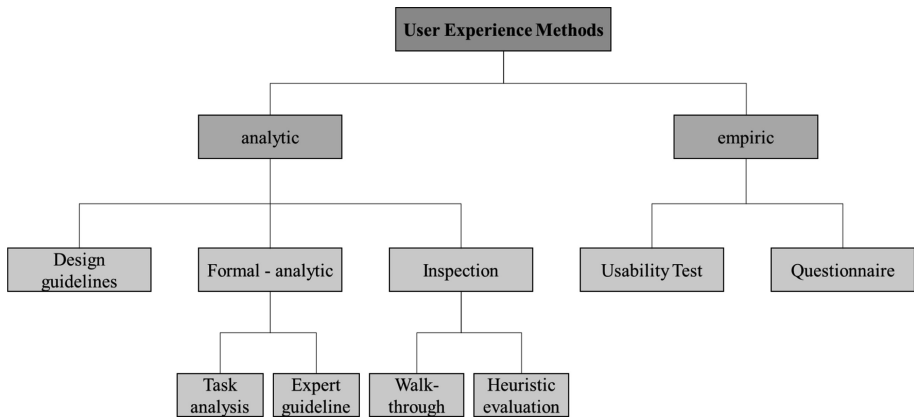


Fig. 1. User experience methods.

The major difference between the shown methods is the evaluating person. In case of analytic methods, there are experts who test and evaluate the certain user experience, whereas empirical methods are based on target-group-specific participants, who verify the system’s or device’s user experience during tests [1].

However, the identification and determination of user irritations is still complex and difficult to achieve. The Customer Experience Tracking Laboratory of the Offenburg University focuses on an innovative method, which is specialized on the emotion measurement, since it allows detailed and purposive conclusions by not exclusively considering usability problems [2].

Customer Experience Tracking (CXT) is a multi-level, modular and scalable process developed by Offenburg University to explore the usability of interactive systems [2]. It is based on the measurement method for user experience in which heuristics, usability tests and interviews are applied. The different methods included within the CXT process are such as analysis of facial expression, eye tracking, think aloud, electro dermal activity analysis as well as motion and gesture tracking, online-questionnaire

Table 1. User experience methods and their characteristics.

	Method	Characteristic
Analytic	Design guidelines	<ul style="list-style-type: none"> • Ensures usability • Easy implementation • List of principles
	Task analysis	• Evaluation of user experience by tasks to be fulfilled by the device or system
	Expert guideline	• Evaluation by rules of software ergonomics (on the basis of checklists)
	Walkthrough	<ul style="list-style-type: none"> • Experts evaluate proposals for the implementation of various features • Verification of an intuitive interface, also for inexperienced users
	Heuristic evaluation	<ul style="list-style-type: none"> • Experts evaluate by using heuristics, for example the ten heuristics according to Nielsen (1994) • Often combined with empirical methods such as usability tests
Empiric	Usability test	<ul style="list-style-type: none"> • Most known method • Target-group-specific participants test the device under real-life conditions by accomplishing tasks
	Questionnaire	• Enables quantitative data

and input reports. The various modules of measurement enable a development of recommendations for action and provide information on customer's emotions during the interaction which are an essential component for user experience [3].

The research project Professional UX, presented in this paper, combines the most common used modules of emotion and user experience measurement to identify signals and irritations caused during the interaction with a device. It includes the three different methods of mimic analysis, voice analysis (received data by using think aloud) and eye tracking, described below.

2.1 Emotion Psychology: Where Do Emotions Come from?

Emotions can be declared as an intermediary interface between an environmental input and a behavioral output that disconnects the stimulus itself from the resulted reaction [4].

Due to changes in state of various organismic subsystems emotions are evolved [2]. They enable the identification of user irritations by applying different methods to measure human's expression and behavior, most importantly the analysis of facial expression and voice [2].

On the basis of the so-called basic emotions it is possible to provide a selection of emotions which are intercultural universally recognized emotions. In accordance with the theory of Ekman and Friesen [6], there are six different basic emotions [5] such as joy, disgust, sadness, anger, surprise and fear with respective changes in facial expressions, which allows the measurement of emotions. The mentioned theory serves as the basis for the CXT- method of Offenburg University.

2.2 Mimic Analysis: How Do Emotions Show?

The facial expression is an important indicator for the existence of an emotion. An emotion is triggered by outer stimuli which cause changes of the person’s facial expression depending on the processed stimuli and triggered feelings. During stationary testings the facial expression is recorded by a webcam. This way the data is available post hoc and can be analyzed and evaluated retrospectively, independently and unnoticed by the participants. The approach of analyzing is by Emotional Facial Action Coding System (EmFACS), a classification for interpreting mimic art. EmFACS is a reduced classification [7] of Paul Ekman’s FACS (Facial Action Coding System) where the mentioned six basic emotions can be related to different action units (Fig. 2) within the definition of the primal scheme of FACS [8]. Due to the 43 muscles in our face an exact interpretation of emotions is possible [9].

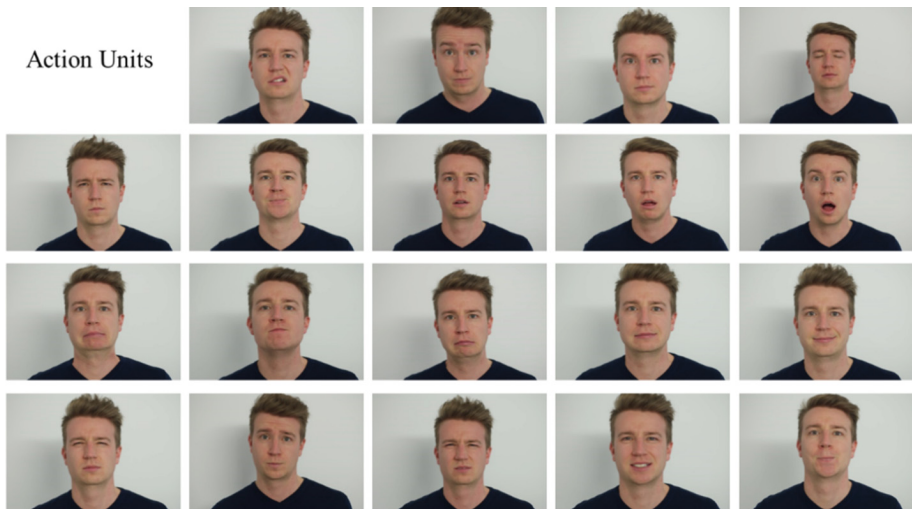


Fig. 2. Action units for the mimic analysis.

2.3 Eye Tracking: What Do Users See?

Eye tracking is one of the method’s central measuring instrument for the determination of irritations and user behavior. The user’s gaze is captured in real time by four cameras and infrared technology. In case of a friction the user’s gaze remains longer at the certain hurdle and is shown as a fixation point. The longer the person looks at a specific spot the bigger the visualization of a fixation point gets. The results obtained are analyzed and interpreted in combination with the mimic analysis and are supplemented by the user’s comments [2] (Fig. 3).



Fig. 3. Eye tracking analysis.

2.4 Voice Analysis: Can We Hear Emotions?

The human voice is not only an audible transmission of words, but rather a doorway into the human inside by providing hidden information and indications. Also, analogous to the changes of facial expression, the voice changes according to an experienced emotion. The way these changes are created and perceived enables the analysis of the acoustic speech signal and backtracking to a certain emotion and thus the origin of a positive or negative event [31].

During user experience testings by means of the customer experience tracking method, there is used the so-called think aloud method. Think aloud provides the opportunity to get information on the participant's thoughts and opinion of the tested system or product. During the tasks every participant has to manage it is requested to comment on positive or negative aspects of the system or device [10]. "Negative affects can make it harder to do even easy tasks" [11] so that especially the negative aspects give some insight into the device's frictions [12]. Think aloud can be run in two different ways. The retrospective think aloud includes a questionnaire and commenting after the testing. Whereas an interview and the participant commenting are conducted during the testing if the concurrent think aloud is used. To sum up, think aloud enables further insight into the user's mind by verbalizing their thoughts and feelings [10].

Thoughts and emotions of users, especially the negative ones, are decisive for the acceptance of devices and for the verification of user experience. For this purpose, the research project Professional UX was initiated to fathom user experience factors and emotions, and to create a device that simplifies the analysis of UX.

3 Professional UX: Soft- and Hardware Selection Criteria

3.1 Innovative Testing Technology: Cost-Effective, Easy to Use, Open Interfaces

Nowadays, since we're facing digital devices daily, it is necessary and has never been more critical to provide frictionless and intuitive experiences. Users expect devices to be optimized and without irritations, which is required as basis [13].

Therefore, it is important to understand expectations and needs of the users as soon and clear as possible and optimize user interfaces by investing in user experience. Within seconds a user is influenced by the system and forms an opinion. It is said, that "\$10 would be spent on the same problem during development, and multiply to \$100 or more if the problem had to be solved after the product's release" [13]. At least 50% of a programmer's time is spent on solving an avoidable problem due to frictions and bad UX [14]. Besides money good UX can save time if problems are fixed at the beginning of development, because they can cause 100 times more time waste during or after the development [15].

Due to the importance of considering user experience in the design process of devices and systems the presented research project is developing a hardware which includes the most common used methods of user experience analysis: mimic analysis, voice analysis and eye tracking. The idea of combining all three modules in a single hardware reduces the costs of purchasing hardware for each of them which would include cameras for the mimic analysis, microphones for the voice analysis and an eye tracker, the most expensive part of a purchase like this (Fig. 4).

Furthermore, the hardware not only allows agencies or UX-specialized companies to perform UX testings, which keeps down costs. Also it allows non-experts, without much expertise, to run tests on their own devices because of an intuitive use. During the development and optimization of the hard- and software within the scope of Professional UX, all usability requirements are considered which leads to a simply structured interface.

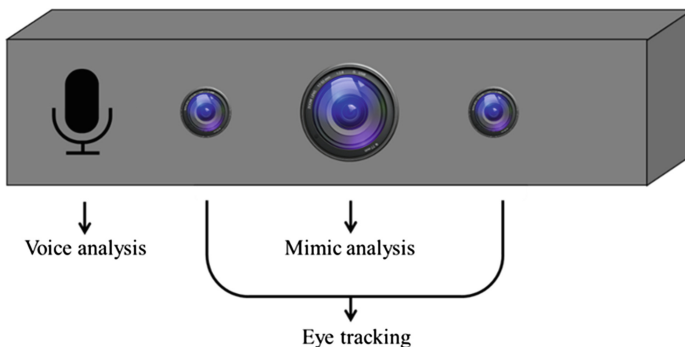


Fig. 4. Professional UX innovation.

The software, inspired by Open Interfaces for mimic analysis, is programmed accordingly so that the user has the possibility of freely selecting the desired module. In addition to this free choice, according to the test's objective, the system starts all selected modules by one single button. Thus a synchrony of all data is given, which is the most difficult and time-consuming part in the evaluation and analysis in the area of UX (Fig. 5).

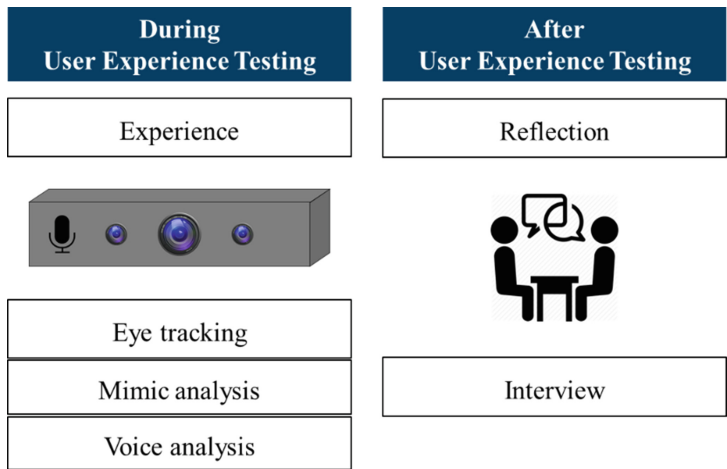


Fig. 5. Modules during and after UX-testings.

3.2 Professional UX: Learning Software

As artificial intelligence we use machine learning methods. Machine learning describes mathematical and statistical techniques that computer systems use to generate knowledge from experience. More precisely, the system automatically learns from training data including features and labels. Features are independent variables that serve as the input in our system and labels are the final output. After the learning process, the system can predict the labels of some test data using their features.

In our application we use machine learning tools to predict the emotion of a user of a graphical interface. On the one hand we analyze the facial expression of the user, on the other hand we make an analysis of its voice. So, we extract both video and audio features from a recording of this user.

We have implemented the machine learning and emotion analysis code in the programming language C++ due to performance reasons. For the facial expression analysis, we make use of OpenCV 4.0 presented in [25]. This is an open source computer vision library including many machine learning tools. For the voice analysis, the Speech Signal Processing Toolkit 3.11 (SPTK), see [26], is utilized to extract numerous audio features. Below we describe our methods in more detail.

The resulting data of our prediction analyses are visualized in a graphical user interface that we have implemented in C#, in Fig. 6.

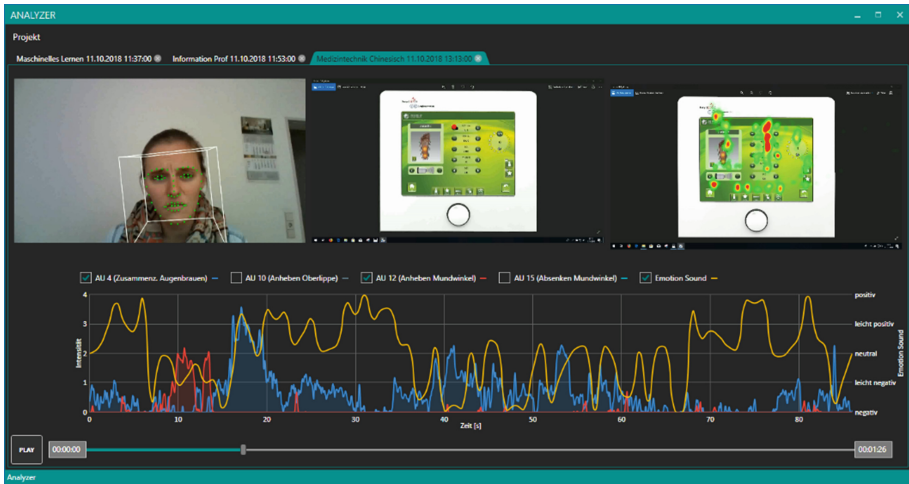


Fig. 6. Professional UX analyzer.

Mimic Analysis

The core of our emotion analysis is the automatic detection of selected action units of the Facial Action Coding System (FACS) by Ekman and Friesen [16].

In the first step, the user face is detected on each frame of the video recording. For this task, we use the very fast method proposed by Viola and Jones [17] implemented in OpenCV where we also take the pre-trained model from. This recognition is optimized respecting our special use case: For example, the face on the frame must have a specific minimum size. We get an accuracy of 99.93% on our test data which does not have challenging head poses, occlusions or lighting conditions.

In the second step, we extract 49 inner facial landmarks from the detected faces. These landmarks are points in the images describing eyes, eyebrows, mouth and nose, and serve later as feature set for the recognition of action units, see Fig. 7. For this landmark localization we use the method by Ren et al. [18] as implemented in (the extra modules of) OpenCV. The training features of this approach are so-called local binary patterns that describe the texture of an image. With these features we trained a model on our own data using random forests as machine learning technique. The normalized mean error of the inter-ocular distance is 6.4% on our test images. The best deep learning methods currently have an error about 4%.

In the third step, the landmarks are normalized. More precisely, we apply an affine transformation to the landmarks such that the distances between the landmarks and the points of a reference face are minimized.

In the fourth step, the goal is to detect the intensities of selected action units of the FACS in order to evaluate the emotion of the user. Seven action units have been selected to be important for emotions relating the user experience of graphical user interfaces (also see Chapter 4 “Study results for mimic analysis”).

For the prediction of the intensities of these action units, we have implemented the algorithm by Werner et al. [19]. The training features of this algorithm are the

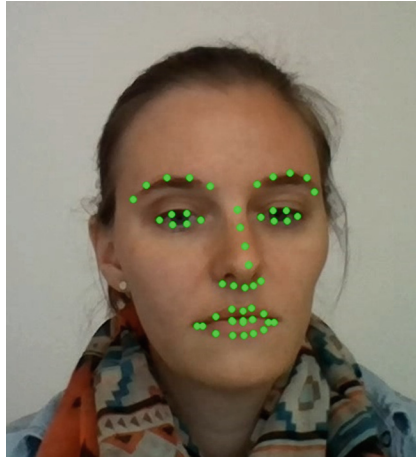


Fig. 7. Landmarks of the mimic analysis software.

normalized landmarks from the third step and the local binary patterns from the second step. The machine learning method is an ensemble of support vector machines with linear kernel. In comparison to deep learning techniques, this procedure is faster with similar detection rate.

In the fifth step, the intensities of the action units are smoothed and calibrated respecting the specific user.

In the sixth step, we predict the intensity of emotion between -1 (negative/irritated) and 1 (positive/not irritated). For this purpose, we use the computed intensities of the action units as training features and take an ensemble of support vector machines with linear kernel as machine learning method.

Voice Analysis

For emotion analysis based on voice we first split the audio recordings into intervals of two seconds. From each of these pieces we extract features by first computing certain acoustic low-level descriptors (LLDs) and then apply statistical functionals to these.

First, we split the audio pieces further into smaller segments (varying between 20 and 60 ms) and sample at 5–10 ms. From each of these samples we extract the low-level descriptors, including pitch, jitter, shimmer and formants as well as mel-frequency cepstral coefficients known to be important in emotion analysis of speech. The SPTK [26] already has methods to extract some of these features which we access in our code. We get a resulting vector containing data from the segments for each LLD and audio piece.

Having this vector, we apply functionals to it like arithmetic mean, standard deviation, arithmetic mean of the slope and some percentiles, resulting in the final feature.

Our feature selection was inspired by the Geneva Minimalistic Acoustic Parameter Set presented in [20] as well as larger known feature sets used in INTERSPEECH

Challenges, for example in the Emotion Challenge 2009 [22] or the Speaker State Challenge 2011 [23]. A nice overview of these feature sets can be found in [21]. We reduced large brute force feature sets by means of recursive feature selection and we selected features which suited our special task.

In the last step we use the features described above to predict the emotion of each interval on a scale between -1 (negative/irritated) and 1 (positive/not irritated). As machine learning method we use a support vector machine with radial kernel which was trained with the extracted features of audio files from the Berlin Database of Emotional Speech, see [24], as well as our own data. On the Berlin Database of Emotional Speech we get an accuracy of 90.2% for the classification negative/neutral/positive with a leave-one-speaker-out cross validation.

4 Professional UX-Solution: Modular System of Intelligent Hard- and Software Components

4.1 Integrated Hardware Components

In order to record synchronized video and audio as well as eye tracking, we have written a PC software to control the recordings. The user can see in advance whether all hardware components are running properly and start and stop recordings, see Fig. 8.

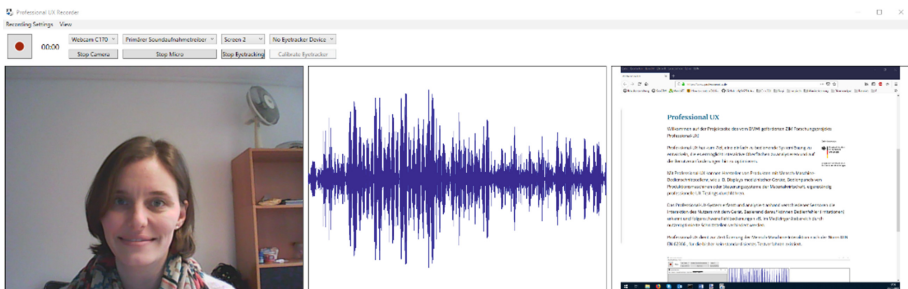


Fig. 8. Recording interface of the professional UX software.

On the hardware site we use a Raspberry Pi 3 B+ as a minicomputer to bundle the recordings. The Raspberry Pi is attached to the computer via LAN or wireless LAN for data transmission and communicates with our recording software via TCP and UDP protocols. To ensure a proper shutdown in case of a premature disconnection we use an uninterruptible power supply, namely a Raspberry USV+ by Ritter Elektronik.

On the recording site we have an Adafruit I2S MEMS Microphone Breakout attached to the GPIOs of the Raspberry Pi. The video recording is done by the Raspberry Pi V1.3 Camera from AZ-Delivery which is connected through the CSI camera port. Furthermore, we use an EyeTech Digital Systems AEye OEM Module for eye tracking via a USB connection.

We specified the recording settings according to our tasks and models from 3.2 as follows: We record video at 25 frames per second and 1280×720 pixel, audio at 44.1 kHz, mono and a sampling depth of 16 Bit. The speed of the eye tracker is 100 Hz and it uses dark pupil methods to calculate the gaze position.

To evaluate the functions of our prototype we conducted empirical studies on Mimic Recognition and Voice Analysis.

4.2 Mimic Recognition Method

For good UX it is indispensable to avoid negative influences caused by frictions. Therefore, a study for the consideration of emotions in user experience was carried out as part of a doctoral thesis at Offenburg University. Measuring UX as the beginning of an optimization process is quite complex, also because of numerous bias effects like social desirability or the experimenter’s effect [2].

Within the doctoral thesis a pilot and a main study were performed with the objective of a mimic-based UX testing method. The method is based on the analysis of the basic emotions for the identification of frictions and aims to set parameters to a few limited trigger points. This way both time and costs can be reduced.

The purpose of the pilot study was to show the existing connection between facial expressions and emotions in the context of interactive applications by clarifying mimic as an emotion-carrier. On the other hand, the main study takes the relation between mimic, emotions and frictions in consideration and aims on finding parameters that help detect irritations. The following table presents the key data of the studies [27] (Table 2).

Table 2. Key data of mimic recognition study.

	Pilot study	Main study
Method	Videoclips which induce specific emotions are implemented in an interactive application and the facial expression is measured by Ekman’s FACS	User irritations that trigger negative emotions are included in an ordering process of an onlineshop and the facial expression is measured by Ekman’s FACS
Sample	33 participants, 363 videos, including 65 min. analysis material	45 participants, 45 videos, 211 min. analysis material

The study’s results identify six different action units for the detection of irritations, in the study related to possible purchase abandonment at an early stage. Initial implementations focus on four following actions units, see Table 3 [27] (Fig. 9).

Concentrating only on a few action units simplifies the distinction between the action units and makes the recognition easy for non-specialists [27].

The application of a mimic-based UX method is conceivable in various areas such as the evaluation of interactive user interfaces of machines, in proactive support of online shops or interactive applications, as well as the support of e-learning applications [27].

Table 3. Relevant action units for irritation identification.

	Action unit	Percentage share	Indicator for possible purchase cancellation
AU 4	Brow Lowerer	85%	✓
AU 10	Upper Lip Raiser	100%	✓
AU 14	Dimpler	88%	✓
AU 24	Lip Pressor	94%	✓

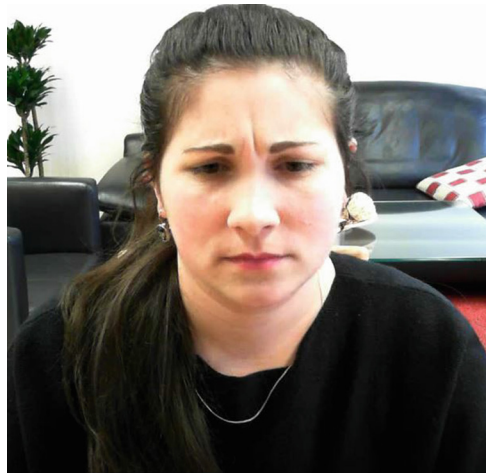


Fig. 9. Example for action unit AU 4.

4.3 Voice Analysis Software

The objective of the presented study on voice analysis is the development of a model that enables the connection of emotions and acoustic changes of the voice. The following hypotheses built the basis of the research within a master thesis at Offenburg University (Table 4):

Table 4. Hypotheses and scientific issues of voice analysis.

No. of hypothesis	Scientific issue
Hypothesis 1	“We change the voice quality depending on the context”
Hypothesis 2	“The voice provides information about the excitement or activation”
Hypothesis 3	“On the basis of voice analysis, positive and negative emotions can be distinguished”

Phonetics describe and examine acoustic speech signals which are generated by a speech apparatus and aim to assign vocal changes to phonation types. Various areas of phonetic allow analysis of speech and voice evaluation. The part of symbol phonetic

“recognizes and holds on to individual sounds and higher linguistic units” by listening carefully and by introspection, whereas phonetics “grasp physiological processes of speech and hearing” and attempts to physically add the acoustic events [28]. The presented study is based on the signal or acoustic phonetic which takes sound vibrations into consideration by the methods oscillogram and spectrogram, see Fig. 10 [29].

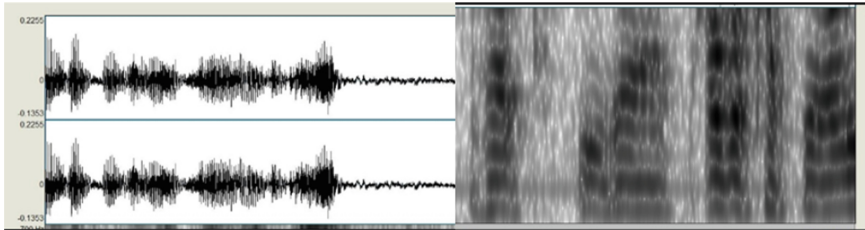


Fig. 10. Illustration of voice analysis data.

The human voice is both an interaction of various muscles and physical reactions (sound waves) for the transmission or the verbalization of information and also an insight into a person’s thoughts and emotions [30].

An acoustic signal changes in response to the emotions which are triggered by certain external stimuli [31]. As an external stimulus is processed and subjectively evaluated, an emotion is triggered which is shown in the way you speak and in the acoustic or sound impression [32]. These changes also include changes in facial expression so that the combination of both methods provide information on emotions and thus possible frictions.

5 Recommendations

Professional UX provides in combination with CXT an easy handling due to a simultaneous start of all modules at the click of a button and the coded automated analysis of all gathered data during the testing.

It simplifies and improves the optimization of interfaces which leads to a reduction of risks and frictions. Using the presented method, which involves the pictured soft- and hardware, the performance of UX testings is made easier and more cost effective due to the combination of the most effective modules for analyzing user experience in a single device and its intuitive use.

In addition, the used methods and modules during Professional UX can also be transferred to other fields like tests of online shops, websites, e-learning platforms or further interactive applications, systems and devices.

Recommendation 1: Listen to your medical user’s needs

Any device is supposed to be useful and lead to the desired objective. Producers of medical devices need to design products with life-saving interfaces. Therefore, it is

indispensable to listen to your customer – both the spoken and the unspoken. Thus medical device providers have to analyze their designed interfaces and go deeper into human minds and emotions to meet their expectations and to ensure a safe usage.

Recommendation 2: Concentrate on simplicity and essentials of interaction elements in medical user-interfaces

The right amount of information is one of the key factors for a successful user experience. Too many options and functions offered during the use can distract and discourage medical users from doing their vitally important tasks. It is essential to support step by step the required tasks and to concentrate on relevant interaction features from the user's angle.

Recommendation 3: Follow the usual - even in specific medical operations

When thinking about a new medical device and during the creative process a positive user experience is fundamental. A big part of it is the warranty of an intuitive use, clarity and reasonable structure. Enabling this kind of medical device interface promotes the interaction and avoids operating errors in, for example, critical surgery situations. One of the most common causes of accidents in medical technology are operating errors that could be prevented by following several UX-guidelines. Many medical devices developed by technical specialists use to be structured in a complex way what leads to unexpected interactions [33] and incorrect operations involving high risks for patients [34]. For this reason, medical devices should be designed intuitively and oriented towards their intended purposes to ensure the patients' well-being.

6 Conclusion

Measuring non-invasive parameters in combination with artificial intelligence methods provides a way to extract quantitative estimations for the user experience of graphical interfaces. Existing methods of machine learning can be extended in a way that basic parameters of the user experience such as stress or confusion during the use of the graphical interface can be determined and visualized.

In particular combining voice and facial expression analysis allows better conclusions in contrast to focusing on just one parameter.

With the combination of the different physical outputs of human interaction, emotion becomes finally measurable. That enables the automatic evaluation of perception results while human-computer interaction happens.

7 Limitations and Future Work

Using complex medical treatment devices is a hazardous challenge: only well-trained persons apply these, but physical conditions of the medical professionals is not always the same. In case of occurring problems there is an urgent need to ensure that no mistake can happen, during a surgery.

The system Professional-UX ensures to provide a user-tested interface for medical device suppliers: all possible issues can be tested and eliminated before the sale of the product starts without huge additional expenses and loss of time.

Nevertheless, there is a chance that new, with the pretesting using the Professional-UX not identifiable problems could occur and endanger the success of the operation. It is necessary to involve medical experts in the evaluation task to make sure that the interface fulfills the needs of the users.

The results of our research project still needs more investigation: The AI-system only learns from usage data and refines with every user testing procedure. A Professional-UX prototype is in use in several companies in southwest Germany. With the gained knowledge of all the different user testing scenarios there will be a huge database for medical device interface design available. This can be helpful for future research projects.

Hochschule Offenburg and Dr. Hornecker Software will enhance their research work and apply for new projects with innovative interaction devices for the medical branch. The challenge is to provide medical experts with the best-fitting interfaces for their life-saving work.

Also, other branches and industries could use the results to optimize processes that involve human machine interaction, like the usage of a display of a forming machine or the usage of an online-shop.

In future work we will also use the recorded eye tracking data for analyzing the emotion. The results of the project provide a good base for the introduction of objective UX assessment methods in the various application areas.

References

1. Sarodnick, F., Brau, H.: Methoden der Usability Evaluation, 2nd edn. s1, pp. 119–191. Verlag Hans Huber (2011)
2. Mueller, A., Gast, O.: Customer experience tracking – Online-Kunden conversionwirksame Erlebnisse bieten durch gezieltes Emotionsmanagement. In: Keuper, F., et al (eds.) Daten-Management und Daten-Services – Next Level, Berlin, pp. 313–343 (2014)
3. Thuring, M., Mahlke, S.: Usability, aesthetics and emotions in human–technology interaction. *Int. J. Psychol.* **42**(4), 253–264 (2007). <https://doi.org/10.1080/00207590701396674>
4. Vogel, I.: Emotionen im Kommunikationskontext. In: Six, U., Gleich, U., Gimmmler, R. (eds.) Kommunikationspsychologie – Medienpsychologie. Lehrbuch (Anwendung Psychologie, 1st edn.), pp. 135–157. Weinheim, Beltz (2007)
5. Schmidt-Atzert, L., et al.: Lehrbuch der Emotionspsychologie. Kohlhammer, Stuttgart (1996)
6. Ekman, P., Friesen, W.: Manual for the Facial Action Coding System, Palo Alto (1977)
7. Noldus: FaceReader: Project Analysis Module. <http://www.noldus.com/facereader/project-analysis-module>. Accessed 21 Oct 2017
8. Paul Ekman Group: Facial Action Coding System (FACS) FAQ - Paul Ekman Group. <http://www.paulekman.com/facs-faq/>. Accessed 21 Oct 2017
9. Argyle, M.: Körpersprache & Kommunikation. Nonverbaler Ausdruck und soziale Interaktion, 10th edn., p. 155. Junfermann, Paderborn (2013)

10. Jo, M.Y., Stautmeister, A.: Don't make me think aloud! – Lautes Denken mit Eye-Tracking auf dem Prüfstand. In: Brau, H., et al. (eds.) *Usability Professionals 2011*, Stuttgart, pp. 172–177 (2011)
11. Norman, D.: Emotion and design: attractive things work better. *Interact. Mag.* **ix**(4), 36–42 (2002)
12. Wong, E.: How to Prevent Negative Emotions in the User Experience of Your Product. <https://www.interaction-design.org/literature/article/how-to-prevent-negative-emotions-in-the-user-experience-of-your-product>. Accessed 31 Jan 2019 (2017)
13. Paunovic, G.: The Bottom Line: Why Good UX Design Means Better Business. <https://www.forbes.com/sites/forbesagencycouncil/2017/03/23/the-bottom-line-why-good-ux-design-means-better-business/#358a58c32396>. Accessed 31 Jan 2019 (2017)
14. Usability.gov: Benefits of User-Centered Design. <https://www.usability.gov/what-and-why/benefits-of-ucd.html>. Accessed 30 Jan 2019
15. Stern, T.: Why User Experience Is the Best ROI Strategy. <https://www.business.com/articles/why-user-experience-is-the-best-roi-strategy/>. Accessed 30 Jan 2019 (2017)
16. Ekman, P., Friesen, W.V.: *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto (1978)
17. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, Kauai, HI, USA, 2001*, p. I-I (2001)
18. Ren, S., et al.: Face alignment at 3000 FPS via regressing local binary features. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014)
19. Werner, P., Saxen, F., Al-Hamadi, A.: Handling data imbalance in automatic facial action intensity estimation. *FERA*, p. 26 (2015)
20. Eyben, F., et al.: The geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Trans. Affect. Comput.* **7**(2), 190–202 (2016)
21. Eyben, F. et al.: Towards a standard set of acoustic features for the processing of emotion in speech. In: *Proceedings of Meetings on Acoustics* (2010)
22. Schuller, B., Steidl, S., Batliner, A.: The INTERSPEECH 2009 emotion challenge. In: *Proceedings of Interspeech, ISCA* (2009)
23. Schuller, B., et al.: The INTERSPEECH 2011 speaker state challenge. In: *Proceedings of Interspeech* (2011)
24. Burkhardt, F., et al.: A database of german emotional speech. In: *Proceedings of Interspeech* (2005)
25. Bradski, G.: The OpenCV library. *Dr. Dobb's J. Softw. Tools* **25**, 120–125 (2000)
26. Imai, S., Kobayashi, T., et al.: Speech Signal Processing Toolkit. <https://sourceforge.net/projects/sp-tk>. Accessed 29 Jan 2019
27. Gast, O.: *User Experience im E-Commerce. Messung von Emotionen bei der Nutzung interaktiver Anwendungen*. Springer Gabler, Wiesbaden (2018). <https://doi.org/10.1007/978-3-658-22484-4>. Dissertation Pädagogische Hochschule Freiburg, ISBN 978-3-658-22484-4, pp. 18–24, 30–46, 131–134
28. Reetz, H.: *Artikulatorische und akustische Phonetik*, 2nd edn., pp. 1–30, 125–137f. WWT Wissenschaftlicher Verlag, Trier (2003)
29. Carstensen, K.-U.: *Computerlinguistik und Sprachtechnologie. Eine Einführung*, 3rd edn., pp. 176–194. Spektrum, Akad. Verlag, Heidelberg (2010)
30. Schwertfeger, B.: Personalauswahl per Sprachtest. In: *Personalmagazin*, no. 12, pp. 32–34. https://www.wiso-net.de/document/PEMA__111524019. Accessed 21 Oct 2017(2015)
31. Schneider-Stickler, B., Bigenzahn, W.: *Stimmdiagnostik. Ein Leitfaden für die Praxis*, 2nd edn., pp. 30–34. Springer, Dordrecht (2013)
32. Schwender, C.: *Medien und Emotionen*, p. 34. Springer Fach-medien, Wiesbaden (2007)

33. Compamed Magazin: Top-Design gegen Pannen im OP. https://www.compamed.de/cgi-bin/md_compamed/lib/pub/tt.cgi/Top-Design_gegen_Pannen_im_OP.html?oid=17100&lang=1&ticket=g_u_e_s_t. Accessed 21 Oct 2017 (2007)
34. Dain, S.: Normal accidents: human error and medical equipment design. *Hear. Surg. Forum* **5**(3) (2002)