

Visual Emotion-Aware Cloud Localization User Experience Framework Based on Mobile Location Services

<https://doi.org/10.3991/ijim.v15i14.20061>

Aiman M. Ayyal Awwad
Tafila Technical University, Tafila, Jordan
AimanAwwad@ttu.edu.jo

Abstract—Recently, the study of emotional recognition models has increased in the human-computer interaction field. With high recognition accuracy of emotions' data, we could get immediate feedback from mobile users, get a better perception of human behavior while interacting with mobile apps, and thus make the user experience design more adaptable and intelligent. The harnessing of emotional recognition in mobile apps can dramatically enhance users' experience. Therefore, this paper proposes a visual emotion-aware cloud localization user experience framework based on mobile location services. An important feature of our proposed framework is to provide a personalized mobile app based on the user's visual emotional changes. The framework captures the emotion-aware data, processes them in the cloud server, and analyzes them for an immediate localization process. The first stage in the framework builds a correlation between the app's default language and the user's visual emotional feedback. In the second stage, the localization model loads the appropriate app's resources and adjusts the screen features based on the real-time user's emotion obtained in the first stage and according to the app's location data collected from the mobile device. Our experiments demonstrate the effectiveness of the proposed framework. The results show that our proposed framework can provide a high-quality application experience in terms of a user's emotional levels and deliver an excellent level of usability that was not possible before.

Keywords—Emotion Recognition, User Experience, Cloud Localization, Mobile Application, Location-Based Services.

1 Introduction

Since the release of mobile platforms, there has been an enormous change from personal computers (PCs) to smartphones [1]. In recent years, the number of mobile and wearable device users has increased rapidly. According to data collected by Statista in September 2020, the number of mobile users worldwide reached 6.95 billion in 2020, with predictions suggesting this is reasonable to increase to 7.1 billion by 2021. In 2024, the number of mobile users worldwide is forecasted to stand at 7.41 billion. [2].

Nowadays, people use mobile and wearable devices at an unusual rate. Their interactions with devices include daily activities, such as making phone calls, sending messages, navigating the Internet, and using different kinds of apps [3]. From smartphones, tablets, smartwatches, and smart glasses, such devices constantly are being developed in capability, performance, and intelligence, which could be utilized to introduce outstanding features [4].

Recently, smartphones are integrated with various physical sensors that can detect a user's action, determine a device's current location, observe physiological signals, capture images, take videos, and record voices [4]. The smartphone's built-in sensors such as cameras, microphones, accelerometers, touch screens, and GPS have been utilized to develop personalized apps and services based on mobile user emotion and current physical location [4, 5]. However, the feedback of such emotion-aware and location-based apps depends on the smartphone itself, which introduces a higher Quality of Experience (QoE) [5].

Although visual emotion-related data and location information can easily be obtained anywhere and anytime by integrated devices (respectively, cameras and GPS sensors), the power consumption and computation complexity of processing and analyzing facial-related data still is a burden on mobile devices. Therefore, a cloud-based emotion-aware app or service is more appealing and required [4].

However, the human's emotion expresses a complex feeling state that occurs due to physical and psychological facial changes. It is typically connected with one's character, mood, spirits, and relations with others. It is essential to read and acknowledge others' emotions as it helps us to react properly and interact effectively in different social circumstances. The ability to precisely recognize users' emotions helps people manage their personal lives and social relationships [4, 5].

Since emotion plays a great role in people's daily lives and social interaction, researchers have been introducing devices and applications that are capable of perceiving, reacting to, and even classifying emotion [4]. Similar to how people discern emotions, devices can interpret a human's mood via people's responses [4, 6]. The facial landmarks and speech have been broadly interpreted as they carry rich information that reflects people's emotions. As people perceive emotive gestures and experience physiological responses, smart devices can also identify human emotions by examining physiological signals [4].

In mobile apps, for instance, the SMS text can be analyzed to recognize the emotion of the users. When the user's emotion is captured, the system can automatically insert an appropriate 'emoji' in the SMS text. Similarly, in the context of emotion, by analyzing a video, a smartphone can automatically adjust the background or play some favorite music according to the end user's emotion [6].

Facial gesture detection is helpful to classify the behavior of a person when a user interacts with the smartphone. Though, a good level of end-user satisfaction or dissatisfaction relies on emotion and sentiment [5]. The challenges concerning users' sentiments or emotions should be deemed to produce an enhanced mobile User Experience (UX) [7].

The UX includes all the end-user interaction aspects with the product in terms of the user's impressions, sentiments, likes, or dislikes that the user has when using

the application [8]. On the other hand, localization is the process of adapting a product to meet the culture, language, and UX of a target locale where the product will be used and sold. Today, in the product development process, UX and localization operate hand-in-hand to satisfy users' goals, mental models, and requirements. Both are central to the product's adoption and success [8, 9].

Localizing the UX needs a strategic approach that goes far beyond text translation; such an approach should create a UX that feels truly local, as it has been designed specifically for language and locale conventions. The localization process enables users to get the best experience in the language of their choice. Truly successful localization requires a user-centered approach to understanding the social and emotional aspects of the intended culture or country [9, 10].

Companies vending mobile apps cannot afford to release apps containing critical defects or UX issues. If users have a poor experience, they will remove the app and turn to a competing product. Therefore, to compete in global markets and expand your business into new markets, mobile app developers need to publish world-ready products. If mobile apps are not localized properly, they will not be used in the local market where the apps' will be sold. Many mobile users have stopped using an app due to lacking localization [9].

Furthermore, the advancement of smartphone computational power and the development of sensors motivate us to take a step further and develop mobile apps that feel what their users feel. A vital aspect of QoE provisioning is to produce a localized UX mobile service based on the user's emotional responses. It is challenging to estimate and quantify the QoE. The sentiment rating feature can be a reasonable way for developers and service providers to measure the QoE based on the emotional states of users collected from their smartphones and then provide a response to improve the localization UX quality [4].

Therefore, in this paper, we propose a cloud localization UX framework that is based on the user's visual emotion and mobile device's geolocation data. We provide an effective solution that addresses a user's negative emotions in a real-time mode. However, to overcome or reduce such negative visual emotion, the proposed framework will instantly localize the app UX design. The localization will create a UX that feels like it has been designed for a certain language and culture. The users should not distinguish that the original product stems from a different country and cultural background. In the proposed framework, we present the system which detects and recognizes facial emotion as well as tells a lot more about that user's mood which could be utilized to get feedback or to know whether a user is satisfied and engaged with the introduced service.

2 Literature Review

The smart features and user acceptability of the emotion-aware mobile apps are the reasons behind the increased growth. Such applications should be in real-time and highly accurate to produce a good level of UX [6]. To the best of our knowledge, this work is the first investigation that combines UX localization, geolocation-related data

for a mobile device, and visual emotion technology in a cloud server. Unlike research works discussed in this section, which only detects user emotions, identifies them, and addresses negative emotions in real-time to deliver a richer UX.

There exist several emotion recognition systems in the literature. The emotion can be realized from speech, image, video, or text [6]. Cloud-based apps are not restricted by storage, processing power, or time. For example, many online game applications capture emotion-aware data, process them in a cloud server, and analyze them for the next enhancement [11]. In this case, the cloud can provide unlimited storage and computation capabilities, and the game developer can have enough time to analyze. Most of the proposed emotion recognition systems do not address all of these issues, as they are developed mainly to work offline and for desktop platforms.

Several emotion-aware applications were proposed in [7] and [11]. Ref. [7] used audio and visual modalities to discern emotion in 5G. They proposed a bimodal system of big data emotion recognition. In their research, a human's speech represents an audio modality, while a human facial image represents a visual modality. They found that these two modalities complement each other in terms of emotional information. Their proposed system delivered excellent speed-up factors (up to 75.55) during the experiments.

Ref. [11] proposed a cloud-gaming framework by embedding the emotion recognition module in it. In Particular, they described a framework that uses audio-visual emotion to adjust the gaming screen with randomly generated effects in order to enhance a gaming experience. Their study showed how a dynamic screen effect, using remote display technology, can lead to emotional responses during gameplay. Also, they showed how the proposed framework stimulates game players and keeps them motivated and engaged. The appropriateness and effectiveness of the proposed framework were estimated using objective and subjective evaluations. Meanwhile, in their research, the negative emotions were addressed and in their future work, more screen effects will be considered to enhance positive emotions.

An intelligent emotion detection system for mobile phones was proposed in Ref. [3]. By using machine learning techniques, the smart keyboard separately interprets a user's emotional states. The proposed system uses accelerometer readings and various aspects of typing style like speed, number of backspaces, and the time delay between letters to train a classifier to predict emotions. Current work has limitations such as classification accuracy, keyboard layout, and registration to the service process.

Ref. [6] proposed an emotion recognition system for mobile apps. In the proposed system, a face image is captured by a smartphone's embedded camera. Some frames are extracted from the image as a representative, and a face detection module is employed to identify the face regions in the frames. The study results showed that the proposed system achieves high recognition accuracy in a reasonable time. In particular, the proposed system has an accuracy of 99.8% using the JAFFE database, and an accuracy of 99.7% using the CK database.

Ref. [5] proposed a framework that presents personalized emotion-aware services by mobile cloud computing and affective computing. With the proposed framework, the traditional mobile cloud computing architecture is adjusted to attain the required QoE in emotion-aware applications. Their goal was to introduce emotion-aware ser-

vices that are personalized, user-centric, and intelligent. However, the proposed system does not take into account the privacy and security issues of user data in the multi-dimensional emotional data acquiring method.

Ref. [12] introduced an initial implementation of a mobile multimodal recognition system for emotion and the creation of an effective database through a mobile app. In the proposed system, the information generated by the sensors of a mobile device is used to perform the recognition of emotions. The recognizer tool runs into a mobile educational application to identify a user's emotions as they interact with the device. The recognition system is developed in such a way that it can be used by different types of mobile apps including educational apps.

A recognition system to discern emotions from facial video on a smartphone in real-time was proposed by Ref. [13]. In the proposed system, the embedded camera of the smartphone is used to capture the face image, BRIEF features are extracted and the k-nearest neighbor algorithm was introduced for the classification. Experimental results showed that the proposed facial expression recognition is successful on a mobile phone and it gives a recognition accuracy of 89.5%. For future work, they intend to enhance further the existing work by adding features such as deep learning for pre-trained facial characters and also investigate other machine learning techniques for classification.

Most of the new educational mobile games introduced for autistic children are not emotion-aware. Therefore, Ref. [1] showed how movable technologies, facial emotion recognition, and voice recognition algorithms can assist autistic kids to learn and enhance their academic skills. In their research, they developed a computerized system, "World of Kids" which promotes the analyses of the emotional evolution of the user while playing a game on mobile devices. They identified the mobile user's emotions via facial image detection to find the best suited and favorable game(s), which may even be a learning game(s).

Some current healthcare systems use emotion recognition to improve patient care. Ref. [14] proposed a method to teach physicians and physicians in training to infer the emotions of the patients to utilize that skill in the context of patient care. Emotion recognition from faces for healthcare was proposed in Ref. [15] and Ref. [16].

3 Globalization, Internationalization, and Localization

When targeting the global market, any company vending mobile apps faces a burning issue to meet all demands of their potential customers. However, if vendors want to expand their business into new markets and serve globally, they have to take into consideration three closely connected processes: Globalization, Internationalization, and Localization [9, 10].

In the first place, globalization is the process of developing and marketing multi-lingual products and services to new customers globally. Globalization is the umbrella term used to describe the internationalization and localization process (see Fig. 1). In essence, globalization means tailoring products or services to a specific geographical

region all over the world by adapting to cultural differences, language, and locale formats [10, 17].

On the other hand, internationalization is usually referred to as the building and designing of a product or a service in such a way that it can be easily adapted to particular local languages and cultures so that it can be marketed easily worldwide without the need for redesign. After the application is internationalized, it can be localized to a particular language and culture. At the design level, internationalization prepares an application for the following phase, termed localization [17].

Particularly, localization is the process of taking a product or a service and making it linguistically and culturally appropriate to the target market (country, language, culture specifics, and desired local “look-and-feel”). Typically, the process includes at least translating all strings of the user interface (UI) textual elements as well as changing the numbering system, changing the date and time formats, currency use, icons, graphics, etc. However, the process provides the appropriate resources for the application based on the device’s locale settings [9, 10].



Fig. 1. Globalization process: Internationalization vs. Localization

Localization and internationalization enable products to reach an entirely new market by adopting them. Internationalization and localization are two essential parts of the software development and design process. Internationalization focuses much more on the designing and engineering process of creating an application, while localization focuses on the customization of content and graphics consumed by local users [10].

For successful localization, developers need a thorough understanding of the cultural context. Otherwise; they will lose the target market. Naturally, languages operate differently. That means there may be issues developers need to consider on the technical side to display text properly. However, the partial app localization might negatively affect UX. Hence, if companies vending mobile apps want to dominate foreign markets, localizing the app is a vital practice towards good UX [9, 17].

When content goes through the localization phase, some of the UX issues concern technical aspects entail customization included [9, 10]:

- Language direction (right to left, left to right, and vertically)
- Flexible localizable UI: How much space is required to display the translated text (may take up less space or more)
- Unicode support: The ASCII character encoding is sufficient for Western European languages' texts. However, languages that use non-Latin alphabets (such as Arabic, Chinese, Korean, and Japanese) require larger character encodings such as Unicode
- Locale preferences (calendar formats)
- Numbering style (e.g., Arabic or Hindi digits)
- Currency format
- System of measurement
- Collation and sorting rules
- Symbols, pictograms, and colors (hand gestures and the use of color to express information)
- Culturally appropriate strings, graphics, references, etc.

In bi-directional (BiDi) language software, text can render in both directions. Arabic and Hebrew, for instance, are written from right to left but numbers are written and read from left to right. The direction of reading and writing affects how information and UI elements should be laid out on the screen (i.e. mirroring awareness) [9, 10].

In the course of localizing apps, icons and symbols are also other issues that may cause problems. Icons show subtle cultural meanings within them. They may have particular negative meanings that may be seen as offensive in the target locale. Symbols that represent hand gestures may be deemed offensive in another region, for example, an OK sign or V-sign may convey different meanings in different cultures [10].

4 Proposed Architecture and System Modeling

4.1 Scenario

Before describing the proposed framework, we present a scenario to show how emotion can impact the UX. Indeed, an application that is not localized correctly and renders resources' contents in a non-native language for end-users may cause some negative emotions like anger or sadness. Assuming that a user is browsing a mobile app, and there is a video camera that continuously captures the user's facial expression and a location sensor that detects a device's current position. The facial expression and detected physical location are sent over the cloud to the localization UX server. Furthermore, consider that the user reaches a behavioral model that makes him/her angry, disgust, or sad. The user may consequently no longer wish to continue using the application. To overcome or reduce such negative feelings, the application's

UX is localized according to the device’s current location. Meanwhile, switching the application’s resources’ contents, locale preferences, layout direction, and visual language must be controlled in a reasonable way to improve overall usability.

4.2 System architecture

The proposed visual emotion-aware cloud localization UX framework spans two levels including the user devices and mobile app. Fig. 2 illustrates the proposed framework. The smartphone with applications collects the emotional and location data of a mobile user sent over the cloud to the localization UX server. The framework focuses on the automatic visual recognition of emotional expressions. However, the proposed method is divided into two different stages: the first one focuses on the detection of small groups of facial landmarks movements and, according to the output, the second decides the associated emotional expression.

Upon recognition of a user’s visual emotion and detection of the smartphone’s geo-location, the emotion recognition API passes a message to the remote smartphone and requests an appropriate localized UX based on the user’s physical location. With the proposed framework, the language switching is triggered according to the predicted user’s visual emotion to achieve the required QoE in emotion-aware services.

4.3 Capturing video and face detection

The facial video is captured by a smartphone integrated video camera. As most of the time, the user faces the screen of the phone; the video mainly captures the face area. Some representative frames are picked from the video. Visual characteristics are extracted from the user’s face. The method locates the facial area in the starting frame of the video sequence and traces the face moves in the succeeding frames. Nowadays, many smartphones have face detection functionality integrated into the mobile system.

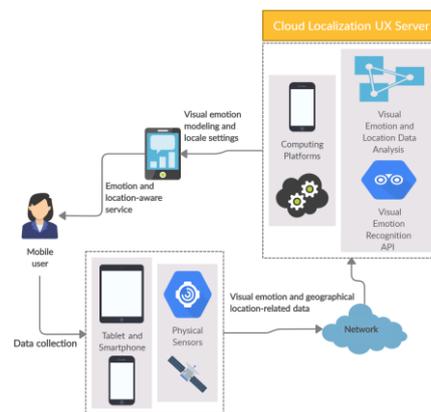


Fig. 2. The proposed visual emotion-aware cloud localization UX framework

4.4 Determining a mobile device's physical location

Nowadays, smartphones are usually equipped with GPS sensors. Due to the many satellites which are orbiting the earth, a GPS sensor can be used for determining the user's location easily. If the GPS signal is weak, the application will fall back to Wi-Fi and mobile network methods. Location-based services provide access to tools that can be employed to determine the device's current physical location. Such location information can be utilized for a wide variety of purposes and enable a smartphone and the application that runs on it to have a better knowledge about its surrounding region and deliver a richer UX [17].

4.5 Visual emotion recognition

The face provides an emotion-rich palette. Humans naturally express and convey their emotions through facial expressions, gestures, and body language [18]. Emotive analytics is an interesting combination of psychology and technology. Thus, many facial expression recognition tools put human emotion into six major categories: Joy, Sadness, Anger, Surprise, Fear, and Disgust. With facial emotion detection, algorithms identify faces within a photo or video and recognize emotion by analyzing the relationship between points on the face, based on selected compiled databases.

Fig. 3 illustrates the process flow of the proposed visual data emotion recognition in mobile devices. Detecting the real-time emotion of the end-user with a camera input is one of the advanced features in the machine learning process. Our proposed technique scientifically detects and reports emotions and facial characteristics using a computer vision library (i.e., OpenCV) and machine learning techniques [19].

In particular, there are special types of Neural Networks - called Convolutional Neural Networks (CNNs) - that are very effective for using images as inputs and can help deliver higher accuracy in emotion recognition [20, 21]. We have used an open-source data set: Face Emotion Recognition (FER) from Kaggle and built a CNN to recognize emotions. The emotions can be classified into five classes: Joy, Surprise, Anger, Disgust, and Sadness.

The primary idea behind the framework presented in this paper is to collect visual emotion data while a user is using a mobile app. In doing so, the mobile camera detects the user's facial expression. Once enough data is collected, preprocessed, and then sent to a cloud server, a Python program is launched in the cloud server to handle the emotional data. Specifically, machine learning techniques are used to build classifiers that can predict the user's current emotional states based on their current visual patterns.

4.6 Localizing the user experience

The mobile app recognizes emotions by analyzing the user's facial landmarks to synthesize the most appropriate UX localized app for the users by loading the most appropriate and favorable resources' contents. Facial gesture detection is helpful to classify the behavior of a person when a user interacts with the app. During app use,

the proposed framework can verify whether the user’s emotional responses and psychological state are positive or negative. If the emotion is negative, it is encoded with 0; otherwise, it is encoded with 1. When the emotion with the negative state is output as the recognized emotion, the zero value is passed as the input to the localization UX process. In the proposed framework, the automatic internationalization and localization approach, proposed in Ref. [17], is used to localize the app according to the location data that the app received from the device.

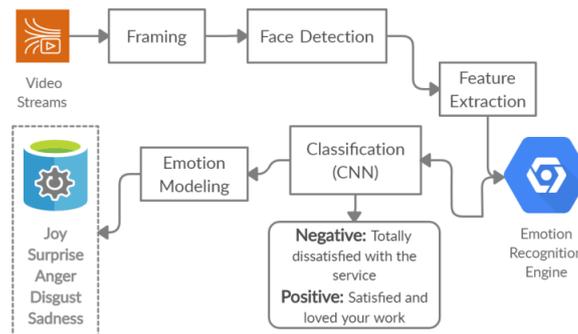


Fig. 3. The visual emotion recognition system

By analyzing a video in the context of emotion, when the user’s negative emotion (i.e., sadness or anger) is detected, the proposed framework checks whether the application talks to the users in their native language. If not, the proposed framework will automatically restart the application with the language spoken in that locale. Moreover, the framework will update the visible UI properly, present a more personal and context-rich UX to overcome such negative visual emotion, and enable users to get the best experience in the language of their choice. After localizing the UX, a user’s negative emotion might be changed to a positive one to keep the user motivated, satisfied, and engaged.

5 Experimental Results and Discussion

To study the feasibility of the proposed emotion-aware cloud localization UX framework, the empirical evaluation was performed with the Pocket Code¹. Pocket Code is an integrated development environment (IDE) for the brick-based visual programming language (Catrobat) for smartphones. Since they do not require previous knowledge of programming syntax, Pocket Code is very popular in programming education for young children, teenagers, and students. However, the Pocket Code is developed to support both the internationalization and localization processes [22].

¹ <https://catrobat.org/>

Fig. 4 presents the actual Pocket Code screens captured from the user’s smartphone. Fig. 4(a) and (c) are the original screens (US English), and Fig. 4(b) and (d) represent the Arabic localized ones. To reduce a negative emotion (a high level of anger or sadness), the proposed feature will load the appropriate localized assets according to the user’s language preferences at the application’s runtime.

To evaluate the quality of the proposed framework, the initial experiments were conducted by inviting 20 high school and undergraduate students in Jordan, namely 16-22 years old. It is valuable to note that the Arabic language is the mother tongue of Jordanians. The students were selected because they have a good knowledge of the block-based visual programming language environment and coding educational tools. The number of selected participants is sufficient since it was stated in Ref. [23] that three to six participants are sufficient to cover 80% of the usability problems. Table 1 shows the demographic characteristics of students with the largest number being females and undergraduates.

Table 1. Demographic characteristics of the participants

Variable	Description	Percentage
Gender	Male	40
	Female	60
Age	16-18 (school student)	30
	19-22 (undergraduate student)	70

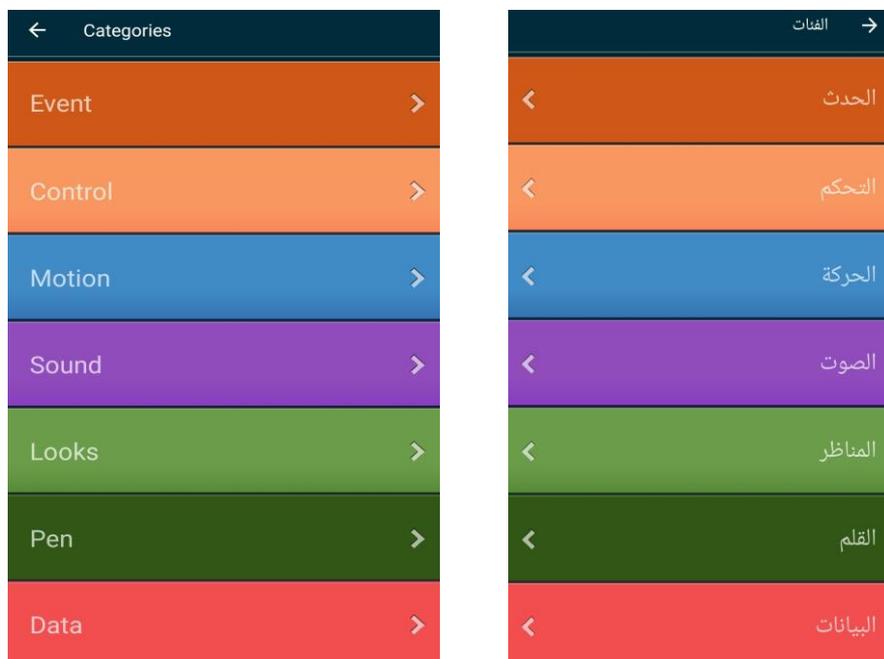
The participants installed and used Pocket Code on their smartphones with a time of 5 minutes. While navigating the original product’s screens (i.e., English language), the various emotional responses of the participants were recorded using the smartphone’s integrated video cameras and an appropriate localization and screen adjustment were provided after processing over the cloud.

The results show that the emotional responses of the 90% of participants are negative (56% are anger, 33% are sadness, 11% are disgust) as shown in Fig. 5, which means that the Pocket Code needs to be localized into the Arabic language. After the localization UX was performed (see Fig. 4(b) and (d)), 88% of those participants are satisfied and engaged with the app.

Satisfaction measures the user’s experience of Pocket Code, their excitement and joy, and the effect of the localization process on users for motivating them to work more. Therefore, after the application’s use, the participants were asked to fill out surveys related to their emotional feedback (e.g., satisfaction, engagement, or disengagement) for the localized application’s items. The survey was created using Google Forms² and distributed through the emails of participants. The responses to the survey were collected, organized, and analyzed automatically. We confirmed that the survey elements were appropriate for achieving our research objectives and are easily understood by the participants in our study.

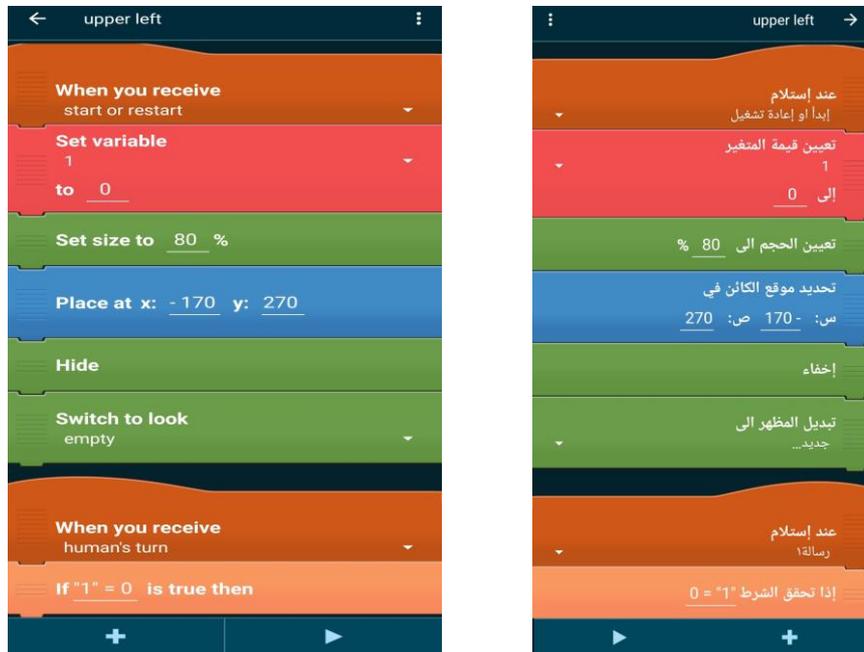
² <https://www.google.com/forms/about/>

To achieve this, the participants were asked to rate each motivational UX item concerning technical aspects of localization (see Table 2) using a 5-point response scale [10]. In particular, they were asked to answer “How satisfied are you with [motivational UX items]?” Mainly, to evaluate the users’ motivation and satisfaction, the totally satisfied, satisfied, neutral, unsatisfied, totally unsatisfied model was adopted. The model was used to evaluate the users’ motivational stimuli in using Pocket Code. The scores of every UX item were calculated based on a 5-point symmetrical Likert Scale to specify the level of motivation and engagement during application use. The Likert scale survey questions are essential in measuring a participant’s opinion towards UX items. It enables respondents to evaluate UX motivational items on a visual scale. A weight is assigned to each icon on the scale to be calculated in the analysis results.



(a) Script categories

(b) Localized script categories



(c) Example of Catrobat’s program (d) Example of localized BiDi program

Fig. 4. Screenshots for Pocket Code

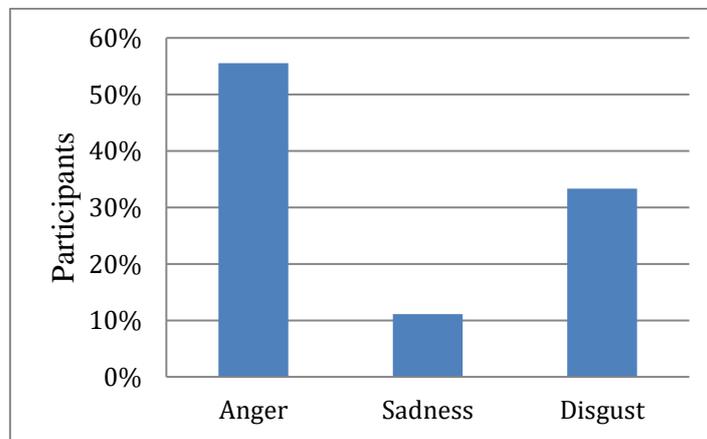


Fig. 5. Percentage of participants negative emotions

The quantitative data were analyzed descriptively using Google Forms. Particularly, we started by collecting and verifying 20 sets of responses for each UX item in the analysis of our questionnaire. Then, we calculated the mean for each item. Table 2 presents the motivational UX items and their respective reported ratings. From the

data analyzed, we found the mean rating values roughly close to 5, which indicate that the appropriate localized screen had some positive effect on users’ emotional feelings during application use, where the highest mean was 4.90 and the lowest was 4.65. Thus, we conclude that our participants were satisfied and engaged with the app.

Table 2. Motivational UX items concern technical aspects of BiDi- languages

Motivational Item	Satisfaction’s Mean Ratings
Bidirectional scripts support	4.70
Flipped design: UI elements render from RTL	4.75
Text direction (Arabic text flow from RTL)	4.75
Encoding scheme: Support any writing script (i.e. Unicode)	4.80
Iconography (such as undo and redo)	4.65
Flexible layout: Translated text may take up more space (text expansion)	4.80
Word wrapping validation	4.80
UI elements alignment (i.e. right-aligned)	4.90
Linguistic requirements: check the semantic correctness of the translated text	4.90
culturally appropriate graphics and colors	4.70

After the localization of the product, the bidirectional languages localization testing methods, proposed by Ref. [10], were performed to test the performance of our proposed framework. The objective is to check the product in terms of complying with Google Material Design guidelines³ in terms of bidirectionality. We ran the Pocket Code (Catrobat’s version for the Android platform) on Samsung Galaxy A51. The testing methods were executed in the Arabic language which was used as the reference. The methods had been created to perform a journey and visit every screen and fragment in the product. They verified whether the expected and actual results are matched or not. However, if the bidirectionality aspects are the same as the expected one, the “Pass” is recorded in the “Result” column; otherwise, “Fail” is recorded in the “Result” column to indicate that the expected and actual aspects are not identical (see Table 3).

However, from Table 3, we can observe that Pocket Code complies with Google Material design guidelines. The Arabic version of the original product looks as if it had been developed in the Arabic user’s local market.

Table 3. Material Design guidelines in terms of bidirectionality

Bidirectionality Aspects	Testing Methods Description	Result
Flipped layout	Text: Arabic text is read from right to left. Timeline indicator: Progress from right to left. Direction-sensitive graphics, like arrows, are flipped. Arabic text is aligned to the right. Icons are displayed on the right of the text field.	Pass
Time progress	UI elements depicting the time progress are displayed in reverse order.	Pass
Localization	Icons that contain text or numbers must be localized.	Pass

³ <https://material.io/design>

6 Conclusion

Visual emotion can play an important role in the mobile app to promote satisfaction and engagement during app use. To sustain a good level of user satisfaction and overcome or decrease such negative visual emotion (i.e., anger, sadness, or disgust), we proposed, in this paper, a framework that employs the user's visual emotion in localizing UX mobile apps which is based on location information. In the proposed framework, a user's visual emotional data is collected by a smartphone. This is achieved by capturing facial video and identifying the face region to determine a user's current emotion. Once the facial features have been extracted, they are given to the emotion classifier. Consequently, if a negative emotion is detected, the application is triggered to be localized according to the device's current location.

The effectiveness of the proposed framework was assessed by empirical evaluations. The empirical evaluation of our framework reported that our approach was able to successfully resolve 88% of detected negative emotions. In a user study, participants rated every motivational localization UX item using a 5-point symmetrical Likert Scale. The results showed that the localization process has a positive impact on the user's negative emotion and provides instant user feedback to alleviate such emotion. However, the current design does not take into account the privacy and security issues of users' data in the visual emotional data collection process. In addition, it would also be interesting to consider different input modalities of emotion (i.e., auditory and visual). These are interesting issues to be investigated in our future work.

7 Acknowledgement

The author would like to thank Professor Wolfgang Slany and his Catrobat development team⁴. We also thank the students who participated in the study. Finally, we would like to point out that the figures in this work were created using Creately App.

8 References

- [1] N. Heni and H. Hamam. (2016). Design of emotional educational system mobile games for autistic children. 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Monastir, Tunisia, pp. 631-637. <https://doi.org/10.1109/atsip.2016.7523168>
- [2] S. O'Dea. (2021). Forecast number of mobile users worldwide from 2020 to 2024. Statista - The Statistics Portal, Statista. [Online]. Available: <https://www.statista.com/statistics/218984/number-of-global-mobile-users-since-2010/> [Accessed Jan. 20, 2021].
- [3] S. Shapsough, A. Hesham, Y. Elkhazaty, I. A. Zualkernan and F. Aloul. (2016). Emotion recognition using mobile phones. IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, Germany, pp. 1-6. <https://doi.org/10.1109/healthcom.2016.7749470>

⁴ <http://developer.catrobat.org/credits>

- [4] J. Shu, M. Chiu and P. Hui. (2019). Emotion Sensing for Mobile Computing. *IEEE Communications Magazine*, vol. 57, no. 11, pp. 84-90. <https://doi.org/10.1109/mcom.001.1800834>
- [5] M. Chen, Y. Zhang, Y. Li, S. Mao and V. C. M. Leung. (2015). EMC: Emotion-aware mobile cloud computing in 5G. *IEEE Network*, vol. 29, no. 2, pp. 32-38, <https://doi.org/10.1109/mnet.2015.7064900>
- [6] M. S. Hossain and G. Muhammad. (2017). An Emotion Recognition System for Mobile Applications. *IEEE Access*, vol. 5, pp. 2281-2287. <https://doi.org/10.1109/access.2017.2672829>
- [7] Hossain M. Shamim, Muhammad Ghulam, Alhamid Mohammed, Song Biao and Al-Mutib Khaled. (2016). Audio-Visual Emotion Recognition Using Big Data Towards 5G. *Mobile Networks and Applications*, vol. 21, no. 5, pp. 753-763. <https://doi.org/10.1007/s11036-016-0685-9>
- [8] Song Wei, Tjondronegoro Dian and Docherty Michael. (2012). Understanding User Experience of Mobile Video: Framework, Measurement, and Optimization. *Mobile Multimedia - User and Technology Perspectives*, INTECH Open Access. <https://doi.org/10.5772/39053>
- [9] Aiman M. A. Awwad, C. Schindler, K. K. Luhana, Z. Ali and B. Spieler. (2017). Improving Pocket Paint Usability via Material Design Compliance and Internationalization & Localization Support on Application Level. *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Vienna, vol. 99, pp.1-8. <https://doi.org/10.1145/3098279.3122142>
- [10] Ayyal Awwad Aiman M and Slany Wolfgang. (2016). Automated Bidirectional Languages Localization Testing for Android Apps with Rich GUI. *Mobile Information Systems*, vol. 2016. <https://doi.org/10.1155/2016/2872067>
- [11] M. S. Hossain, G. Muhammad, B. Song, M. M. Hassan, A. Alelaiwi and A. Alamri. (2015). Audio-Visual Emotion-Aware Cloud Gaming Framework. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 12, pp. 2105-2118. <https://doi.org/10.1109/tcsvt.2015.2444731>
- [12] M. L. Barron-Estrada, R. Zatarain-Cabada and C. G. Aispuro-Gallegos. (2018). Multimodal Recognition of Emotions with Application to Mobile Learning. *IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, Mumbai, India, pp. 416-418. <https://doi.org/10.1109/icalt.2018.00104>
- [13] H. Alshamsi, H. Meng and M. Li. (2016). Real time facial expression recognition app development on mobile phones. *12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, Changsha, China, pp. 1750-1755. <https://doi.org/10.1109/fskd.2016.7603442>
- [14] Ragsdale John, Deusen Reed, Rubio Doris and Spagnoletti Carla. (2016). Recognizing Patients' Emotions: Teaching Health Care Providers to Interpret Facial Expressions. *Academic Medicine: Journal of the Association of American Medical Colleges*, vol. 91, no. 9. <https://doi.org/10.1097/acm.0000000000001163>
- [15] S. Tivatansakul, M. Ohkura, S. Puangpontip and T. Achalakul. (2014). Emotional healthcare system: Emotion detection by facial expressions using Japanese database. *6th Computer Science and Electronic Engineering Conference (CEEC)*, Colchester, England, pp. 41-46. <https://doi.org/10.1109/ceec.2014.6958552>
- [16] Zhang Yin, Chen Min, Huang Dijiang, Wu Di and Li, Yong. (2016). IDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization. *Future Generation Computer Systems*, vol. 66, no. 0167-739X, pp. 30-35. <https://doi.org/10.1016/j.future.2015.12.001>
- [17] Aiman M. Ayyal Awwad and Nur EL-Din EL- Rez. (2018). Automatic Internationalization and Localization Based on Android Location Services. *International Journal of Computer Science Issues (IJCSI)*, vol. 15, no. 5, pp.42-51. doi:[10.5281/zenodo.1467655](https://doi.org/10.5281/zenodo.1467655)

- [18] S. M. Pablos, J. G. García-Bermejo, E. ZalamaCasanova and J. López. (2015). Dynamic Facial Emotion Recognition Oriented to HCI Applications. *Interacting with Computers*, vol. 27, no. 2, pp. 99-119. <https://doi.org/10.1093/iwc/iwt057>
- [19] Akhmad Qashlim, Basri Basri, Haeruddin Haeruddin, Ardan Ardan, Inggrid Nurtanio, Amil Ahmad Ilham. (2020). Smartphone Technology Applications for Milkfish Image Segmentation Using OpenCV Library. *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 14, no. 8, pp. 150-163. <https://doi.org/10.3991/ijim.v14i08.12423>
- [20] Yulius Harjoseputro, Yonathan Dri Handarkho, Heronimus Tresy Renata Adie. (2019). The Javanese Letters Classifier with Mobile Client-Server Architecture and Convolution Neural Network Method. *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 13, no. 12, pp. 67-80. <https://doi.org/10.3991/ijim.v13i12.11492>
- [21] Amir Dirin, Nicolas Delbiaggio, Janne Kauttonen. (2020). Comparisons of Facial Recognition Algorithms Through a Case Study Application. *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 14, no. 14, pp. 121-133. <https://doi.org/10.3991/ijim.v14i14.14997>
- [22] K. K. Luhana, C. Schindler and W. Slany. (2018). Streamlining mobile app deployment with Jenkins and Fastlane in the case of Catrobat's pocket code. *IEEE International Conference on Innovative Research and Development (ICIRD)*, Bangkok, Thailand, pp. 1-6. <https://doi.org/10.1109/icird.2018.8376296>
- [23] Jakob Nielsen. (1994). Estimating the number of subjects needed for a thinking aloud test. *International journal of human-computer studies*, vol. 41, no. 3, pp.385–397. <https://doi.org/10.1006/ijhc.1994.1065>

9 Author

Aiman Mamdouh Ayyal Awwad is an Assistant Professor of computer science. He obtained his Ph.D. in computer science from Graz University of Technology/Austria in 2017 with research interests related to mobile computing. Currently, he has intensively engaged with various administrative roles at Tafila Technical University such as heading the department of Computer Science and heading of the Computer and Information Technology Center Council starting from mid of 2020. He is also a reviewer in several IEEE international conferences. He has more than 9 publications in various international journals and conferences. His research interests include mobile computing and applications, digital image processing, machine learning, and cybersecurity. Email: AimanAwwad@ttu.edu.jo

Article submitted 2020-11-25. Resubmitted 2021-02-07. Final acceptance 2021-02-09. Final version published as submitted by the authors.