December 11, 2018

Hand of the Invisible Poet

Anonymous

Hand of the Invisible Poet

Abstract

The present disclosure describes a method for automatically superimposing content of a user's preference on targeted portions of images captured by the user in real time. The targeted portions to display the content, also known as secondary spaces, are detected via heat maps, which are based on the user's focus on the screen. The secondary spaces are regions, where the user's attention is least centered, that is, the regions, where the color density is least. Moreover, the method provides the user with the content depending upon the user's location, so that the user can relate and make use of the generated content in the present scenario.

Introduction to the problem statement

Images shared globally, especially the content-based images tend to contain a lot of free space. This eventually goes unnoticed to the users viewing the images. Choosing a designated space in an image to add content such as quotes, poems, etc., requires a lot of the user's effort in positioning it, so that the content will not overshadow the main areas of the image. Currently, there are many webpages and applications, which offer a variety of motivational, funny and optimistic content to the user but unfortunately, such content is independent of the user's location. Also, the driver behind looking for such content is the user. It is not only cumbersome to search the required content but also finding the most appropriate content considering the user's present location is arduous, because of availability of eclectic content. Hence, the user's judgment for choosing an appropriate content is perplexing.

The techniques described here offers the content to the user without any recurring inputs from the user but is automatically provided to the user basis the user's location and his/her mood. For instance, if the user is present in the office, the present system will automatically detect the user's location and will provide the content which will uplift and motivate the user to work proficiently in the office. Initially, the system will ask for the user's preference regarding the poem, poets, themes or specifically the content, which he/she would prefer to be shown to him/her basis his/her mood and location.

Displaying the content to the user is also a primary concern since it superimposes the images, which the user has captured over a period or at specific location, such that the actual image will not be clouded by the content. Therefore, the present invention tracks the eyes of the user to analyze spaces on the captured image, where the user is more focused. The present invention also analyzes those spaces, which are not under the user's focus. These spaces can then later be utilized to superimpose the content and subsequently, display to the user. The system provides freedom to the user to share it on the social media. Thereafter, the system takes a feedback from the other users to categorize if the generated content is appropriate or not and if it is appropriate, how likeable it is among the other users.

Description of the disclosed system

The user interacts with the present system, which consists of an AI engine and other sensors such as eye tracking sensors, which detects spaces, where the user is more focused. Moreover, a gaze sensor is used to first detect the user's attention towards the screen of the system and later the same gaze sensor triggers the eye tracking sensor to initiate the tracking mode. The gaze sensors and eye tracking sensors can be embedded into the camera of the system. These sensors also work as mood tracker of the user, that is, the sensors track the facial expressions of the user at certain instances. The system comprises of a database, which stores a reference dataset of facial expressions. The data of the user's facial expression, as fetched via sensors is then compared to the reference data set, so that the actual mood of the user can be detected. Further, an interface is provided to the user where he/she can manually choose and select different options, that can be offered to the user at intervals, such as quotations, factual data etc. The system comprises a micro-processor that works with a set of instructions, which processes inputs from the user and the database, and then generates an output basis the set of instructions. AI engine includes a set of instructions embedded in the micro-processor. The AI engine learns and improves performance of overall system. The AI engine learns from the user's feedback as well as from the feedback received from the other users. There are certain additional sensors, which further ensure complete user interaction with the system, such as a global positioning system and gyroscope.

Further, an on-screen selection mechanism on the system allows the user to provide input to enable or disable the working of the system. The present invention can be implemented on, but is not restricted to smartphones, laptops, computers, tablets, etc. Figure 1 depicts the components of the present system.

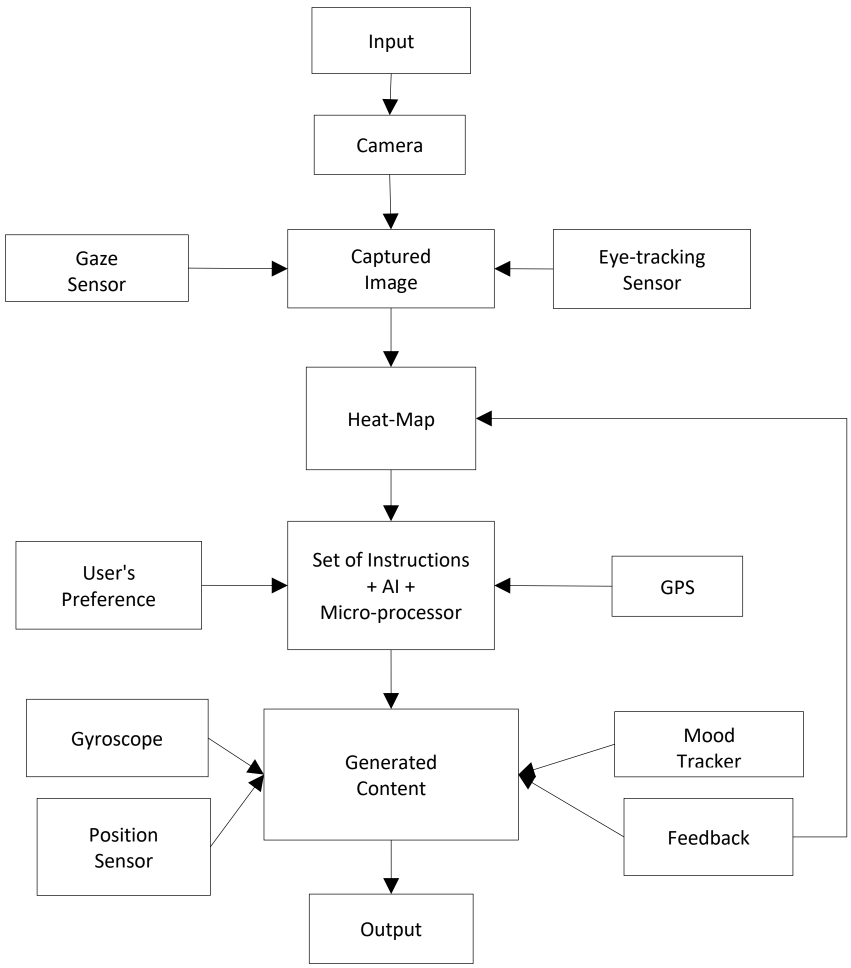

Figure 1 - Components of the present system

Working

A user selects an application installed on the system, which allows the user to capture images. The captured image is displayed to the user. The user is provided with an option for content creation, which when chosen initiates a process of an invisible poet. Meanwhile, the gaze sensors detect if the user is gazing on the screen of the system or not. If the gaze sensor detects the user's attention towards the system screen, it triggers the eye tracking sensors to track the user's retina while looking at the system screen. Eye tracking sensor feeds the real time data directly to the micro-processor, which processes the data and generates a heat map. The heat map highlights spaces on the screen, where the user looks the most and where the least. The spaces, where the user sees the least is then marked as secondary spaces, where the content is to be displayed. The generated heat map is then stored in the database.

After the secondary spaces are detected and marked, a global positioning system detects the user's current location and records it as one of the parameters for sorting content to be displayed. For example, if the user is present in the office, the system would sort content that are inclined towards work-life balance or productivity quotes. Initially, the user's preferences such as poets, themes etc., are already stored in the database of the system. The preferences can further be amended as per the user's request. Inputs from the mood tracker also affect the content that is to be displayed to the user. Mood tracking can be facilitated via sensors, which are embedded into the system. Once the user captures an image, the sensors will track the facial expressions of the user and based on the comparison of the reference dataset stored in the database and the user's tracked facial expressions, the system will detect if the user is sad, jolly, tensed, nervous etc. Depending on the user's mood, identified via the comparison, the system will display only that content, which will enhance the user's mood. If the user is sad or nervous, the system will showcase the content, which rejuvenates the user's mood. Feedbacks from previous posts on social media as well as from the user also play a pivotal role in decision making in the system. For example, if the user has marked down a previous post, the system will learn and generate a better content from next time.

Once the system has finalized the content that is to be superimposed on the captured image, the set of instructions will load the content on the secondary spaces. To facilitate this, the captured image as well as the content are stored into the database at separate instances. The micro-processor will then analyze the heat map stored in the database. The micro-processor will superimpose the content on the captured image, based on the analysis of the heat map. After integration of the captured image and the content, the result is then displayed to the user as a final image. As soon as the user focuses his/her attention on the final image, the sensors track reaction of the user after seeing the final image. Data set of user's reactions is then stored into the database. This data will serve as the feedback for the next content creation.

The next step allows the user to share the final image on the social media. On the system implementing the above-mentioned steps, that is, capturing image, adding content and displaying the content to the user, an additional option would be available for sharing. Once the user selects the option, the system would allow the user to share the final image on any of the social media as per the user's preference as shown in Figure 2.

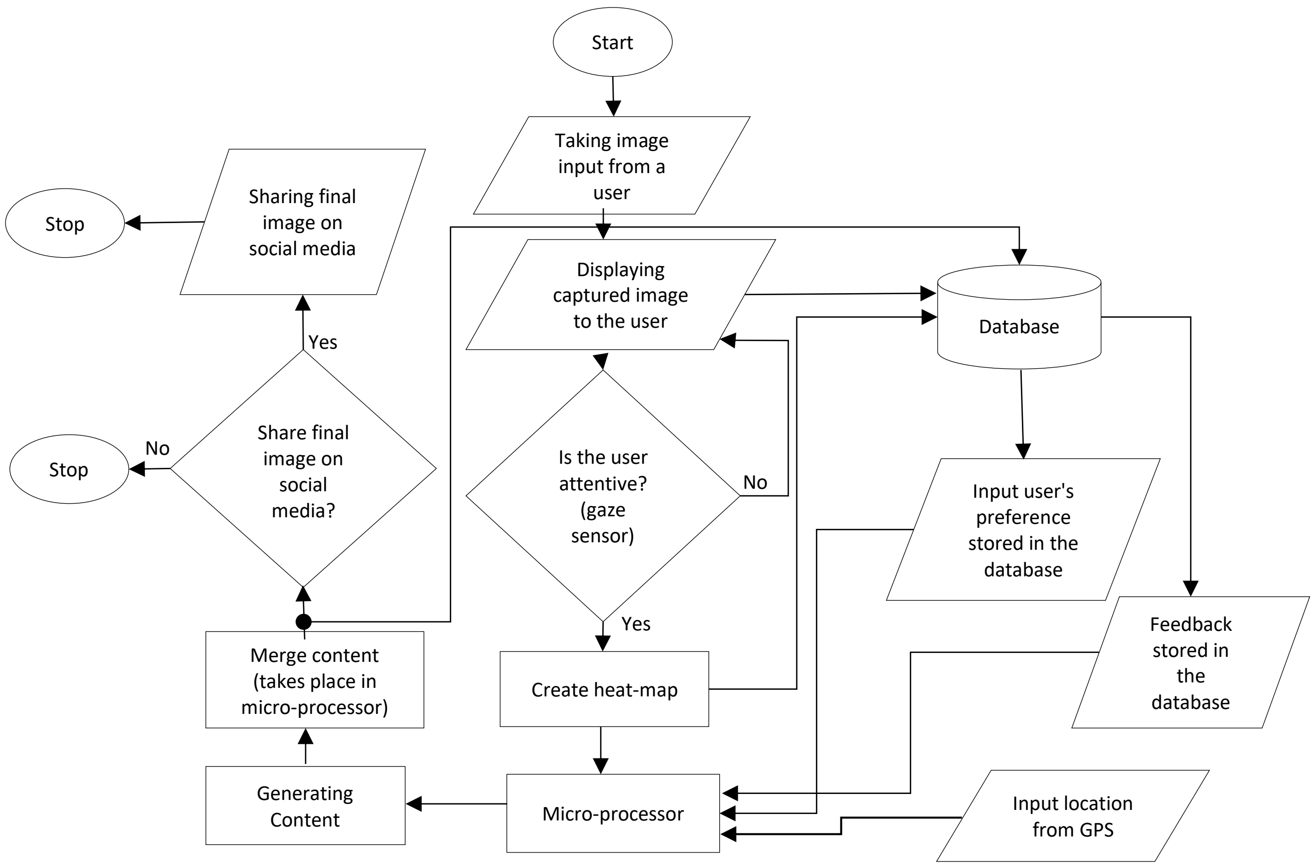

Figure 2 - Flow diagram of the present invention disclosure

The system can be further extended for implementation in an Augmented Reality (AR) device. The content that is generated by analyzing the user's mood and location is superimposed on secondary spaces of real-time visuals displayed on a screen of the AR device.

Extended Features

1. Background Relevancy – This embodiment ensures detecting presence of any background information on the captured image and providing content related to the background of the captured image. For example, if the user is capturing an image in front of the Statue of Liberty, the sensors embedded in the system will automatically detect the location of the user and extract the background of the captured image and later identify any building, monument, garden or any place of importance.

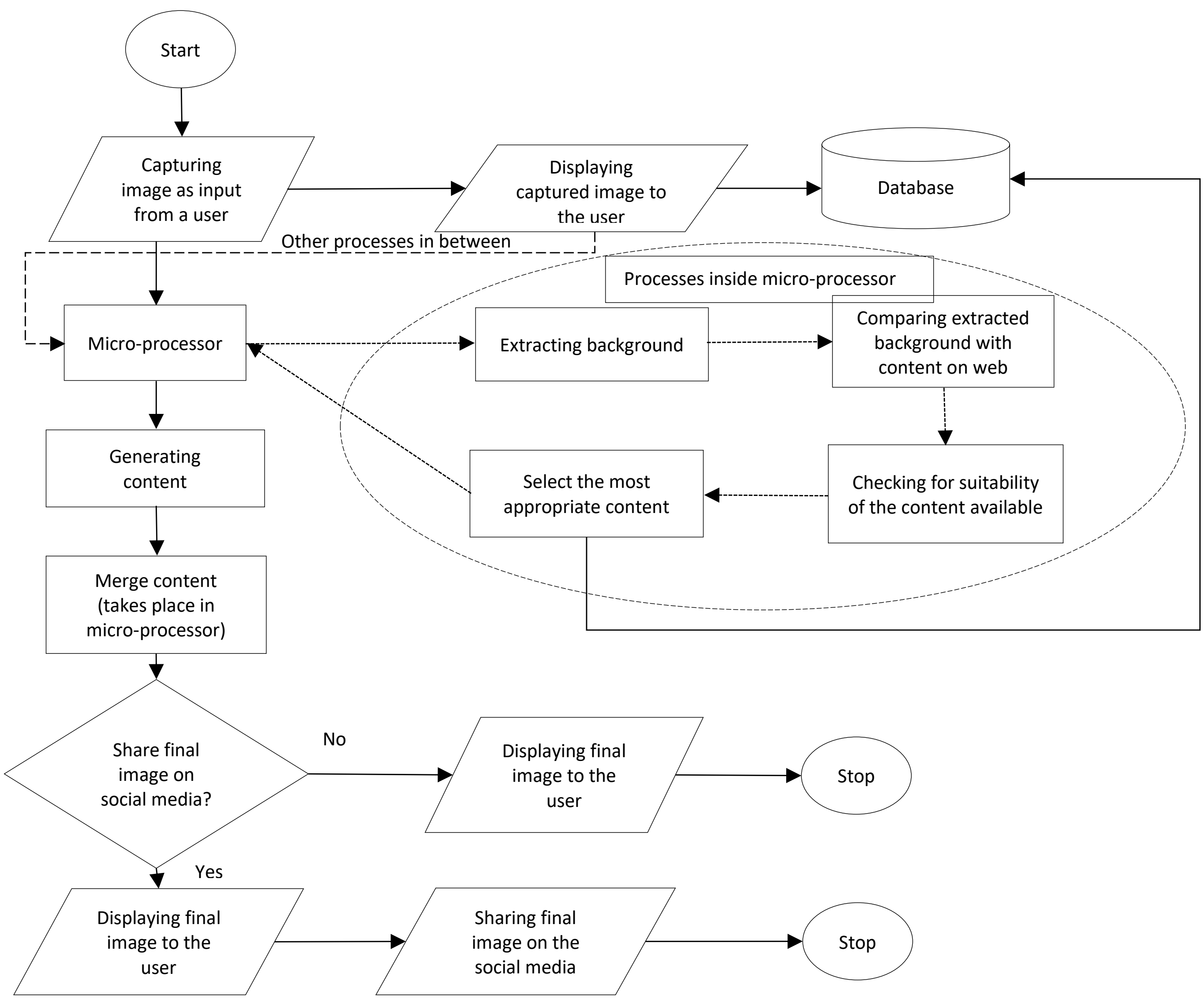

Figure 3 - Flowchart of additional embodiment checking and displaying content of the background in captured image

For identification, the system will extract images from the web to compare the extracted images from the web and of the background of the captured images. Based on comparison, the system will identify the background of the captured image. Basis the result of identification, the system will search for appropriate content related to the extracted background from the captured image and store it in the database. Before storing, the system will look for flags such as likes, emoticons etc., present on social media platforms, to find how likeable the content is among the other users. Basis the likeability factor, the system will make a call if the content is appropriate to be appended to the captured image or not. Again, based on the heat map, the system will superimpose the content on the captured image where the user's attention is least pronounced, as summarized in Figure 3.

2. Online review – In the previous embodiment the system involves taking feedback from other users, which will eventually influence decision making on the type of the content to be displayed to the user. The current embodiment facilitates the above-mentioned operation by using likeability scores. After the final image with superimposed content is generated, the user can share it on a social media platform. On sharing, the system will keep track of the reaction of the other users to the shared content. The other users will mark the content as good, excellent or bad based on different inputs, which are not limited to the likes or emoticons. A value is pre-assigned to each emoticon and stored in the database against the emoticon. Such that emoticons, which represents an "excellent" reaction have a score of 2, emoticons for a "good" reaction have a score of 1, whereas emoticons representing a "dislike" from users have a score of 0. After a certain period, the system will take an average of the reaction scores from the other users and provide them to the micro-processor, which will serve as an input for the feedback loop and thus, become a parameter for next content generation as shown in the Figure 4.

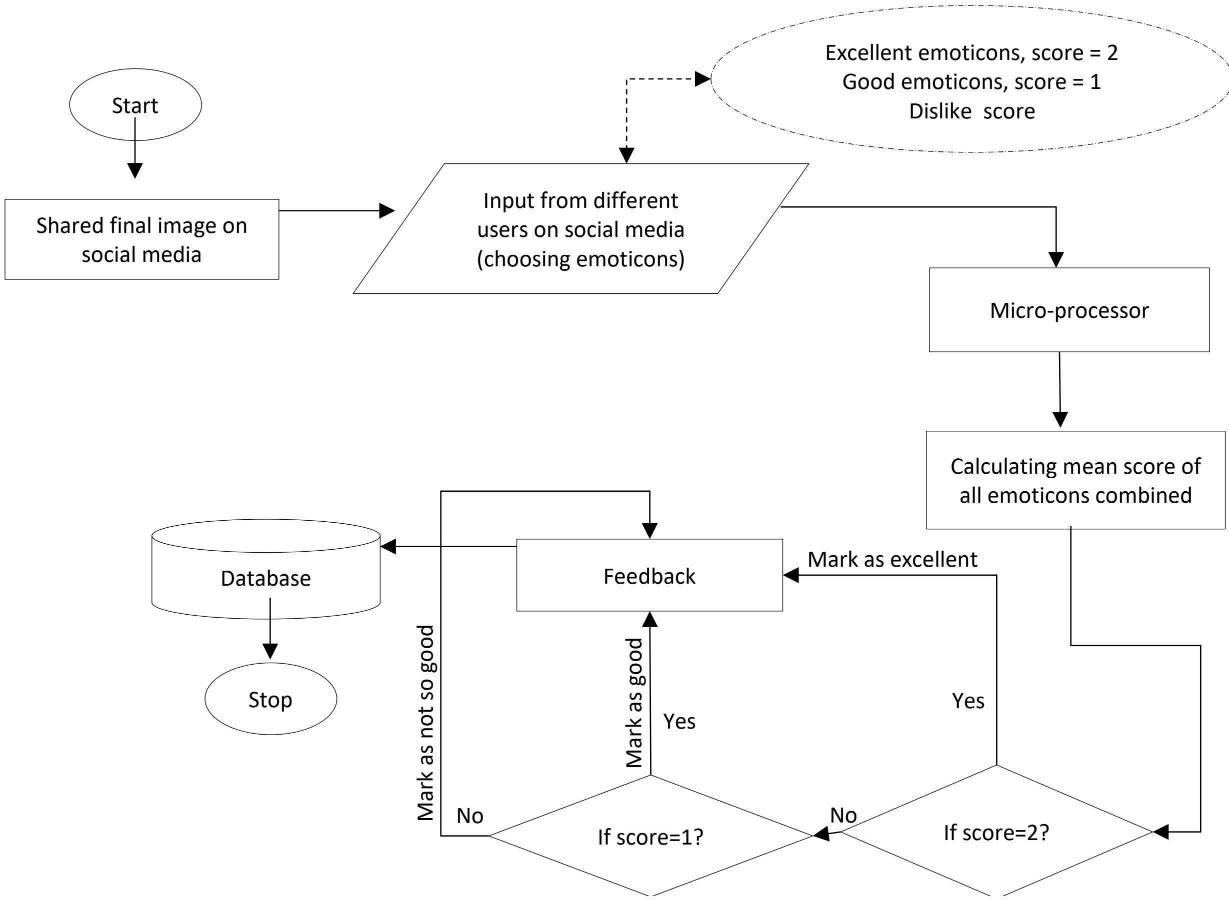

Figure 4 – Flowchart for generating feedback of the generated content via reaction of viewer's online

3. Gesture-based analysis – In yet another embodiment, the system will facilitate tracking of the user's bodily reaction to analyze if the final image appears satisfactory to the user or not. The system consists of sensors such as gyroscope, camera and gaze sensors, which record the user's motion when the final image is displayed to the user. On seeing the final image, if the user makes gestures such as head tilt, head thrust or head retreat, the system will track and analyze such gestures and thus, help in providing better feedback. The system is facilitated with pre-defined bodily gestures that are tagged against a likeability score and stored in the database. Gestures like head tilt, which showcase intermediate user interest have the likability score of 1. The gestures which showcase least interest, such as head jerk, shoulder shrug etc., have the likability score of 0. Moreover, the gestures which showcase high interest towards the displayed final image such as forward node etc., have a likability score of 2. These pre-defined gestures will serve as reference gestures for future analysis. Additionally, the system is enabled with sensors, which will track the user's bodily gestures and then provide the tracked data (referred to as G_i) into the

micro-processor. The micro-processor will simultaneously take input from the database, where pre-defined bodily gestures (referred to as G_d) are stored and compare the user's tracked bodily gestures with the pre-defined bodily gestures.

Once the gestures are identified, a likeability score is obtained from the database corresponding to the input bodily gesture, which will serve as the feedback for the final image displayed to the user. This quantitative feedback will also take part in decision making for enhancing the next series of outputs as shown in Figure 5.

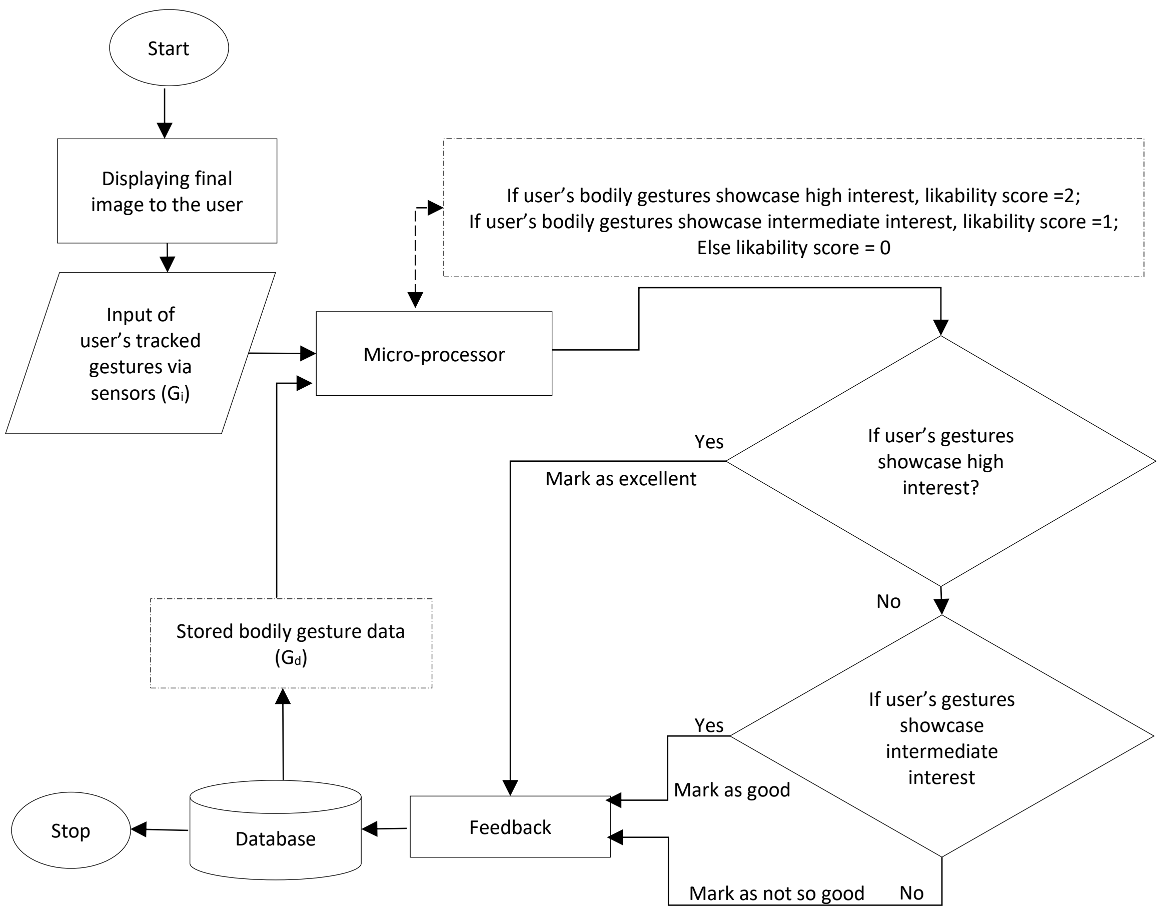

Figure 5 – Flowchart for generating gesture-controlled feedback on the generated content

The present invention offers a unique way of superimposing content on captured images without any user intervention. Moreover, the system takes all necessary inputs automatically, such as user's location and user's content preference. The superimposition of the content on the capture image solves a major challenge of saving user's time and effort. Furthermore, the described method of taking feedback from the user also enhances the user-machine interaction.