Chang, M. et al. (Eds.) (2019). Proceedings of the 27th International Conference on Computers in Education. Taiwan: Asia-Pacific Society for Computers in Education

The Development of a Hand-Washing Education System

Katsumi NAGATA[a], Masaki OONO[b] & Masami SHISHIBORI[b*]
[a]*Affiliation A, NTT Plala Inc., Japan*
[b]*Affiliation B, Tokushima University, Japan*
*bori@is.tokushima-u.ac.jp

Abstract: This study develops a system that focuses on proper hand-washing patterns to examine whether hands are being washed properly. By comparing images of hand-washing taken at sinks and the like with model images, it informs users in real-time whether they are using proper hand-washing patterns. Optical flow and skin-color pixel areas are used as feature amounts for recognizing proper hand-washing patterns. SVM (Support Vector Machine) is introduced as the recognition model in this system.

Keywords: health education, image processing, machine learning, support vector machine

1. Introduction

Every year, the likes of influenza and norovirus become problems to the extent that Japan's Ministry of Health, Labor, and Welfare releases information regarding their spread and relevant vaccines and treatments. They are highly contagious diseases with intense symptoms that can easily become severe.

Gargling and hand-washing are ways to prevent their spread. Particular importance is attached to hand-washing. However, few people are cognizant of proper hand-washing. Causes of this include not knowing how to properly wash one's hands, hand-washing being perceived as a job, and it being difficult to see whether one's hands have become clean.

This study developed a system that focuses on proper hand-washing movements to examine whether hands are being properly washed. By comparing videos of hand-washing taken at sinks and the like with model hand-washing videos, it informs users in real-time whether they are engaging in a proper hand-washing pattern.

The outline of this system is shown in Figure 1. The learner washes the hand in actual environment and records the movement of hands by the fixed camera. And then, hand-washing video is uploaded to the system in real time. The system analyzes the video, and informs the insufficient hand-washing patterns to the learner.

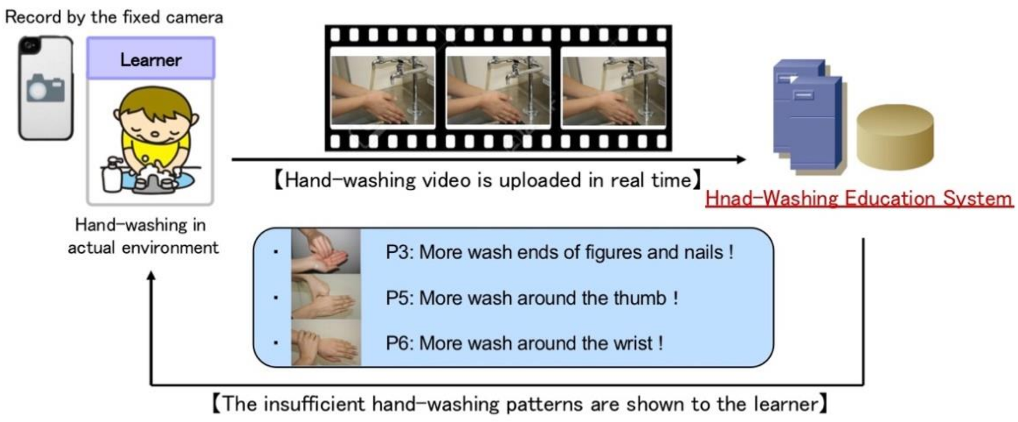

Figure 1. Hand-Washing Education System.

2. Related Works

Much research has been carried out on hand-washing, and it has been proven that proper hand-washing is effective from a hygiene perspective. However, existing examination methods are not easy to carry out. For example, they involve kits that use chemicals or specialist observation.

In existing scholarship, an examination system for proper hand-washing has been developed that uses movement and form characteristics related to hand gestures found in sign-language recognition (Igari, & Fukumura, 2016) and gesture recognition (Sawada, Hashimoto, & Matsushita, 1998). However, there has been the issue of decreased recognition accuracy due to noise included because of hand-shaking during washing and changes in lighting. Therefore, when removing causes of noise and capturing movement characteristics, the decrease in accuracy that results from hand-shaking is addressed by this study by carrying out correction that takes into account the hands' center of gravity.

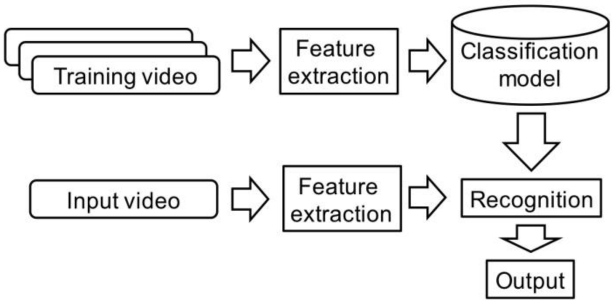

Figure 2. System Flow.

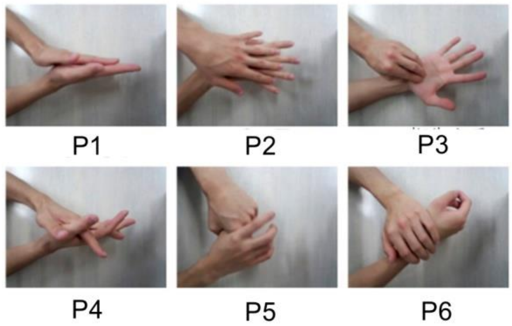

Figure 3. Proper Hand-Washing Method.

3. Hand-Washing Education System

3.1 System Construction

Figure 2 shows the flow of hand-washing examination system. For the input video, hand-washing videos shot in real-time are used. For the training video, proper hand-washing videos are used. We used SVM (Suykens, & Vandewalle, 1999) as a recognition classifier.

The proper hand-washing method is, as shown in Figure 3, "P1: Scrub palms," "P2: Scrub back of hands," "P3: Scrub ends of fingers and nails," "P4: Scrub between fingers," "P5: Twist opposite hand around thumb," and "P6: Rotate wrist in palm of other hand." After examining whether the six patterns have been followed, the output shows undetected patterns.

3.2 Feature Extraction

3.2.1 Skin-Color Image Extraction

Using HSL color spaces, the skin-color region is cut out as the hand region. By carrying out contraction and expansion on this skin-color region, noise is removed. Moreover, based on labeling, noise is removed so that only the largest area in the image is left behind as shown in Figure 4.

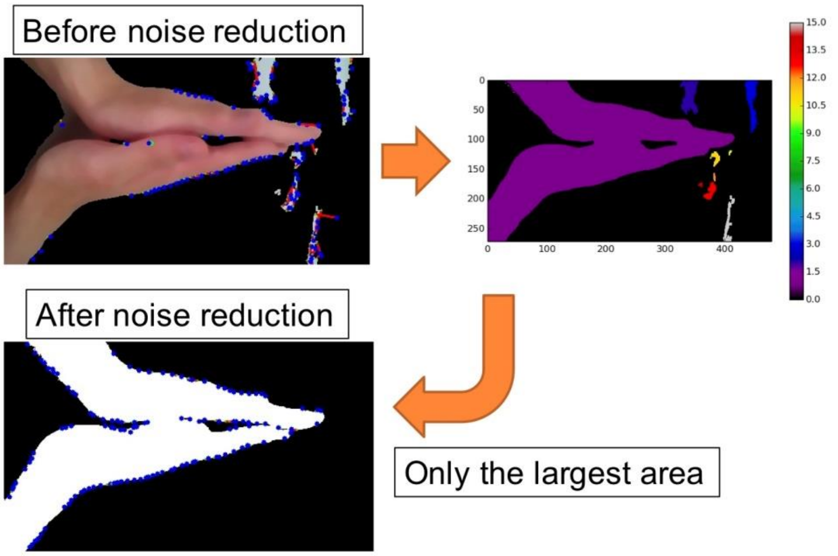

Figure 4. Example of Noise Removal Image.

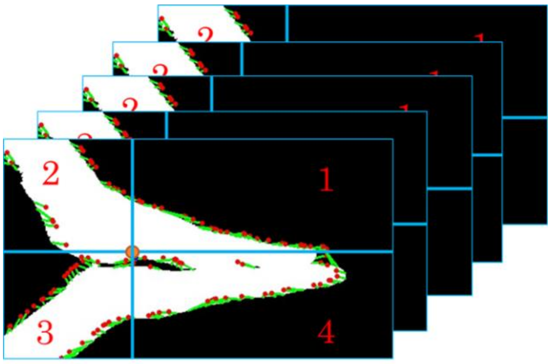

Figure 5. Example of Skin-Color Area.

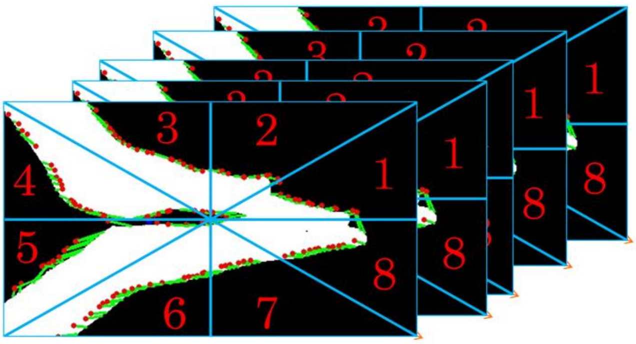

Figure 6. Optical Flow Acquisition (8 Categories).

3.2.2 Skin-Color Area Extraction

Feature extraction is done by, as shown in Figure 5, identifying the hands' center of gravity and dividing the skin-color area into four ranges (1: 0°–90°，2: 90°–180°，3: 180°–270°，4: 270°–360°). The skin-color areas in the divided ranges are captured every few frames, and the average area values for each range are used as feature amounts.

3.2.3 Optical Flow Extraction

The optical flow shows the pixel movement vectors from the previous frame to the current frame. The pyramid LK method, which iteratively implements the LK method (Lucas-Kanade method) (Lucas, & Kanada, 1981), was used as the optical flow detection method. This method withstands noise well and involves little computational effort, enabling stable and swift detection. It is therefore well-fit for a real-time examination system.

Feature extraction was done by calculating vector angle T acquired in the optical flow (Figure 6), and categorizing angle T into eight ranges (1: 0°–45°, 2: 45°–90°, 3: 90°–135°, 4: 135°–180°, 5: 180°–225°, 6: 225°–270°, 7: 270°–315°, 8: 315°–360°). Every few frames, the appearance frequency and size of the vectors in the categorized ranges are acquired. The vectors in each of the ranges are used as appearance frequency and the average size and variance as feature amounts.

4. Evaluation

4.1 Experimental Environment

In order to investigate the effectiveness of this method's feature extraction, our experiments only recorded the movement of hands in order to remove as much as possible the external factors that could get in the way of the features. External factors for hand-washing include soap bubbles, water from the faucet, the environment around the sink, and so on. Therefore, for both the input and training videos, we placed a camera at a fixed height on a table and filmed multiple people washing their hands, directly from the above. A premise of our experiment was that proper hand-washing involves doing the six patterns for 30 or more seconds (each pattern for five or more seconds).

For our database, with the six patterns in one person's hand-washing video (one pattern lasting approximately five seconds) as a set (six videos), we filmed eleven sets for six people. In total, we prepared 66 sets (396 videos). On each video, we carried out feature extraction and created a categorization model. We used C++ (OpenCV v. 2.4.9) for our development system language. For our development equipment, we used a web-camera to acquire images in real-time.

Table 1 *Accuracy Using the Only Skin-Color Area*

Input / Output	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*
P1	*16%*	*0%*	*25%*	*50%*	*0%*	*8%*
P2	*0%*	*58%*	*33%*	*8%*	*0%*	*0%*
P3	*0%*	*8%*	*83%*	*8%*	*0%*	*0%*
P4	*16%*	*0%*	*0%*	*83%*	*0%*	*0%*
P5	*0%*	*0%*	*0%*	*0%*	*91%*	*8%*
P6	*8%*	*16%*	*33%*	*0%*	*16%*	*25%*

Table 2 *Accuracy Using the Only Optical Flow*

Input / Output	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*
P1	*66%*	*0%*	*0%*	*33%*	*0%*	*0%*
P2	*0%*	*100%*	*0%*	*0%*	*0%*	*0%*
P3	*0%*	*16%*	*75%*	*8%*	*0%*	*0%*
P4	*8%*	*0%*	*0%*	*83%*	*0%*	*8%*
P5	*0%*	*0%*	*0%*	*0%*	*100%*	*0%*
P6	*0%*	*0%*	*0%*	*0%*	*8%*	*91%*

Table 3 *Accuracy Using Both the Skin-Color Area and the Optical Flow*

Input / Output	*P1*	*P2*	*P3*	*P4*	*P5*	*P6*
P1	*58%*	*0%*	*0%*	*41%*	*0%*	*0%*
P2	*0%*	*100%*	*0%*	*0%*	*0%*	*0%*
P3	*0%*	*16%*	*83%*	*0%*	*0%*	*0%*
P4	*0%*	*0%*	*0%*	*100%*	*0%*	*0%*
P5	*0%*	*0%*	*0%*	*0%*	*100%*	*0%*
P6	*0%*	*0%*	*0%*	*0%*	*0%*	*100%*

4.2 Evaluation: Is Each Pattern Properly Recognized?

We carried out experiments to determine if each pattern is properly categorized. To do so, we used one set (six videos) out of the 66 sets (396 videos) as input data and the remaining 65 sets (390 videos) as training data. Via cross-validation, by switching the input data set by set with the training data, we carried out a total of 66 experiments. Three types of methods were used to output data in order to compare each feature: the only skip-color area, the only optical flow and both of them.

We have shown the results of the experiments using these methods in Tables 1 through 3. The vertical axis shows the input and the horizontal axis the results of recognition. The closer the results of the diagonal axis to 100%, the higher the categorization accuracy.

Using the results in Tables 1 to 3 to compare the overall proper recognition rate of three proposed features, the average correct recognition rate of the only skip-color area was 59.3%, and the average correct recognition rate of the only optical flow 85.8%, and the average correct recognition rate of both the skip-color area and the optical flow was 90.2%. Our proposed two features have improved accuracy complementary, showing a higher correct recognition rate.

From Table 3, we can find that the category P1 has a lower correct recognition than other categories. Especially, about 40% of input data P1 are recognized to the category P4 by mistake. This is probably caused by hand-washing styles of P1 and P4 being very similar. Because not only the hand shape but also the hand movement direction of their hand-washing styles are very similar, P1 and P4 can't be classified by the only skin-color and optical flow. In order to classify between P1 and P4, new features other than the skin-color and optical flow will be required. Moreover, it appears that in order to increase overall recognition accuracy, measures need to be taken for videos with low accuracy in which skin color is not being properly detected (for example, Figure 7).

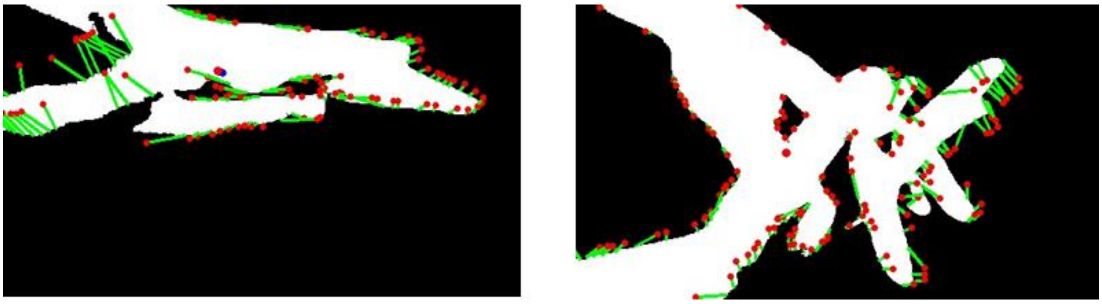

Figure 7. Cause of Decrease in Accuracy (Example of Failed Skin-Color Detection).

5. Conclusion

This study, comparing a method found in existing scholarship and our proposed method, carried out experiments to test feature amount effectiveness. Our experiment results showed that recognition of each pattern is possible, confirming the effectiveness of our confirmed method's feature amounts.

Future tasks include examining videos for which our method was not very accurate, improving upon the causes of this, as well as carrying out experiments using the feature amounts regarding external factors such as soap bubbles and water.

References

Igari, S., & Fukumura, N. (2016). Recognition of Japanese Sign Language Words Represented by Both Arms Using Multi-Stream HMMs. *Proceedings of the 7-th International Multi-Conference on Complexity, Informatics and Cybernetics*, 157-162.

Sawada, H., Hashimoto, S. & Matsushita, T. (1998). A Study of Gesture Recognition Based on Motion and Hand Figure Primitives and Its Application to Sign Language Recognition. *Transactions of Information Processing Society of Japan*, *39(*5), 1325-1333.

Suykens, J. A. K. & Vandewalle, J. (1999). Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*, 9, 293-3000.

Lucas, B. D. & Kanada, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision. *Proceedings of the 7-th International Joint Conference on Artificial Intelligence*, 2, 674-679.