Simple and Practical Skin Detection with Static RGB-Color Lookup Tables: A Visualization-based Study

Frode Eika Sandnes[1,2], Levent Neyse[3] and Yo-Ping Huang[4]
[1]Faculty of Technology, Art and Design, Oslo and Akershus University College of Applied Sciences, Oslo, Norway
[2]Faculty of Technology, Westerdals School of Art, Communication and Technology, Oslo, Norway
[3]Kiel Institute for the World Economy, Kiel, Germany
[4]Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan 10608
Frode-Eika.Sandnes@hioa.no, levent.neyse@ifw-kiel.de, yphuang@ntut.edu.tw

Abstract— **Many skin detection approaches have been proposed in the image analysis literature. Some are simple and static; the others are dynamic and rely on complex machine learning algorithms and training data. Generally the simple approaches are preferred. We hypothesize that the developers' choice for the simple approaches are due to the reasonable quality of results and ease of implementation, since the results of more sophisticated results are not readily available. This paper explores the skin color of a large number of hand samples using color space visualization. The results suggest that a static method may suffice for many applications, but that a small set of rules is not enough to capture the details of skin. Moreover, the results suggest that successful skin detection does not depend on the color space used as there are no apparent advantages of using a perceptual uniform color space such as CIElab. A skin detection approach based on RGB-color table-lookup is proposed that is able to capture the complex skin color cluster shape. The method is practical and simple to implement with minimal computational cost. The lookup table is released into the public domain.**

Keywords- skin detection, color analysis, clustering, image analysis

I. INTRODUCTION

Skin detection is used within many application domains including the tracking of people using their faces [1, 2], hand gesture recognition [3], biometric authentication such as fingerprint and palm print recognition, and digit ratio measurements [4, 5, 6].

The studies on skin detection is abounding [7, 8]. The literature seems to have grown drastically with the emergence of the first low-cost digital cameras in the mid-90s and the research field has been active until the present day. Many approaches have been explored such as simple static rule-based methods [9, 10], rules in different color spaces [11], dynamic rule based methods based on histograms [12, 13] to more sophisticated dynamic methods that employ various complex algorithm such as Gaussian mixed models [14] and principle component analysis [3]. Researchers have also explored image preprocessing to enhance skin detection [15]. The computational effort involved in skin detection has also been explored [16] and optimizations suitable for real-time operation have been proposed.

The problem of skin detection involves segmenting an image into skin regions and non-skin regions. In its general form nothing should be assumed about the background. The problem is relatively easy to solve if the difference between the background and the skin is large. For example, many of the hand images used in this study captured using flatbed scanners all have dark unsaturated backgrounds that are highly distinct from the more saturated skin color pixels. However, even with controlled environments such as flatbed scanners there are problems. Some scans may show the scanner bulb behind the hands leading to ambiguously colored pixels. Alternatively, the scanner background may have a similar color to the skin.

This study involves the classification of pixels as skin or non-skin, thus the detection of, say, a hand on top of the background of other skin is beyond the scope of this study. To solve such problems, a two-stage process may be useful, that is, first to detect the skin regions, and next to detect the foreground hand from the background skin using spatial features.

Many of the skin detection studies adopt dynamic approaches on the assumption that the images are highly variable. This study is based on the position that the skin color remains relatively fixed irrespective of different skin types, and therefore that static methods may be suitable for many applications. Although backgrounds may be highly variable and complex, it is not that relevant if one only need to search for the presence of colors in the skin part of a color space.

Despite that many sophisticated and allegedly robust skin detection methods have been proposed paper citation counts reveal that many researchers prefer the simple rule-based methods despite their obvious limitations. One reason for the popularity of the simple methods could be that they are relatively easy to understand and quick to implement, while at the same time producing acceptable results. Moreover, the researchers behind the more sophisticated methods seem to have failed to disseminate successfully skin-data or code that can be easily reused by other researchers. Researchers who come across the need for skin detection do not necessarily have access to large skin databases to replicate these researchers work. One objective of this study is also therefore to provide other researchers with a concrete skin detection model that can be deployed directly out of the box.

Some of the skin detection research has also considered the choice of color space [17, 18] with the argument that a color space more similar to the human vision system is preferable. Perceptible uniform spaces such as CIElab are argued to be very close to the human visual system, but these models are not widely understood by developers and they involve intricate computations. This study also addresses color models and it is argued that the choice of color space is less important, while the simplicity and understandability of the models are more important. Skin detection in the RGB-domain may be preferable as it is well-understood and have no significant effect on the results with a lookup-table approach.

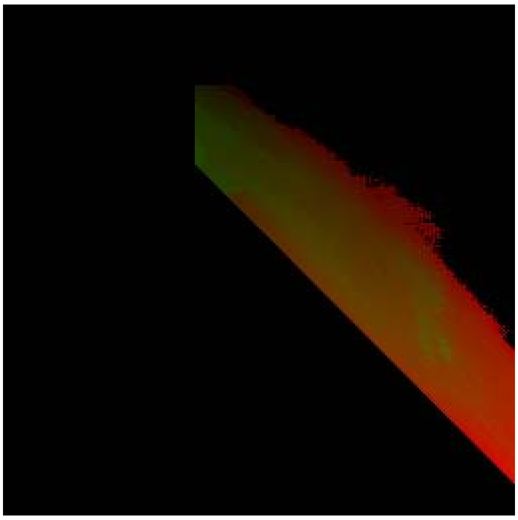

Figure 1. The 256 × 256 skin-color lookup-table in run length encoded format. X and Y coordinates correspond to the red and green components respectively. The red component indicate the offset in the blue dimension and green the length of the skin cluster along the blue dimension. Black indicate that there is no skin color for a given red-green value pair.

II. METHOD

The method proposed herein involves a volumetric lookup table. The table is implemented as a 256 × 256 × 256 Boolean array where the truth values determine if a given RGB color vector is skin or not. The memory requirements of the lookup table are unproblematic with current hardware. Once the lookup table is built, each lookup pose very limited computational effort. The complete lookup table is presented in Fig. 1 as a 22.1 Kb image with the following representation. The x and y coordinates correspond to the red and green values of a RGB color vector, while the red and green components of the image represent the start of the skin volume along the blue component and the length of the skin volume along the blue component. This representation is effectively a simple form of volumetric run length encoding [19]. The image blue component is not used. The lookup table and corresponding code is released into the public domain. The following sections outline the construction of the lookup table.

A. Source data

A database of high quality images of hands acquired using flatbed scanners from 286 individuals with different demographics such as gender, geographic origin, etc., were collected. These scans were acquired in different labs by different researchers with different equipment. Most of the images had a resolution of approximately 2500 × 3500 pixels.

The images were manually classified into easy cases and difficult cases, resulting in a total of 174 easy cases. Easy cases were defined as scans where the hands are very distinct from the background with either a very dark or green background as skin color is believed to contain very little green. Difficult cases were defined as images with a background with similar hues to the hand. The rationale for sorting the images into easy and difficult images is that the easy images could be classified using one of the simple rule-based methods from the literature.

B. Peer's rule based model

A highly cited rule-based method [9] classifies a pixel given by the red, green and blue vector R, G, B as skin if the following conditions are satisfied:

$$R > 95 \qquad (1)$$
$$G > 40 \qquad (2)$$
$$B > 20 \qquad (3)$$
$$\max(R, G, B) - \min(R, G, B) > 15 \qquad (4)$$
$$|R - G| > 15 \qquad (5)$$
$$R > G \qquad (6)$$
$$R > B \qquad (7)$$

This approach is simple to implement and efficient to compute. The fact that the model is defined in terms of the RGB model makes it attractive to programmers. This method is capable of successfully classifying most skin pixels in our set of 174 easy cases.

Next, it is common to post-process images that have been binarized to eliminate noise. For this purpose morphological operators are common, such as a morphological closing operators with a full 5 × 5 structural element.

C. Analysis of Peer's rules

Fig. 2 shows a visualization of the simple color model based on the simple rules defined by Peer et al. [9]. The visualizations were generated using the open source tool CloudCompare. The top two images show two views of the skin model in the RGB color space while the two bottom illustrations show two views of the skin model in the HSV space. These visualizations reveal the obvious fact that the rules are over simplistic. Another observation is that the space defined by the model has simple geometric shapes in both color spaces.

The regions defined as skin in Peer's model are relatively large. It is therefore probable that Peer's model will lead to a high false positive rate as many non-skin pixels will be incorrectly classified as skin. This observation is confirmed when trying to classify our difficult cases using this model as it is unsuccessful in separating the hand from the background.

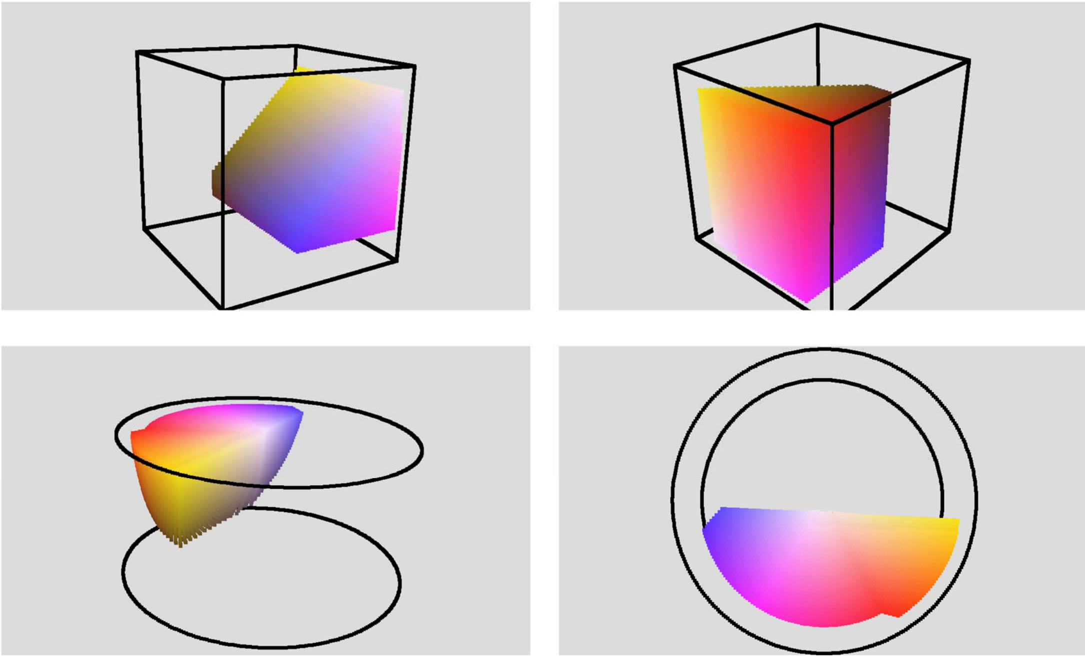

Figure 2. A visualization of the simple skin color model. Top views: RGB-space, bottom views, HSV-space.

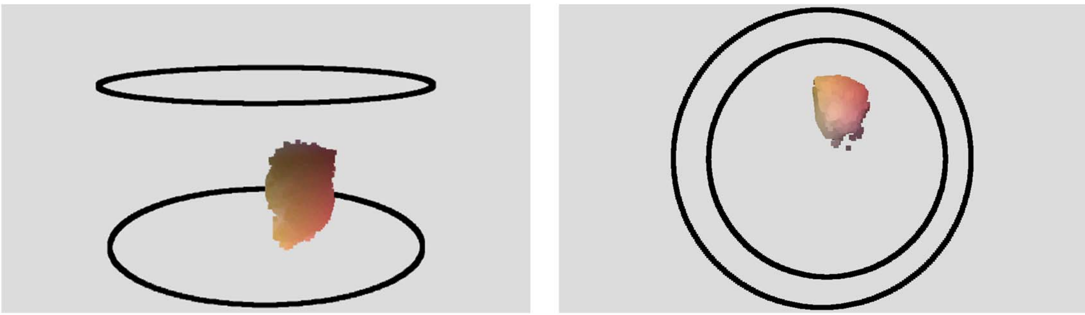

Figure 3. The visualization of the skin pixels for one single individual in HSV-space. Left: side view, right: top view.

The assumption of this study is that the simple rule-based model includes most skin pixels as skin pixels, but that a more realistic model will have a smaller volume. Fig. 3 confirms this assumption as it shows two visualizations of the skin pixels for one single individual. Clearly, the pixels all lie within the bounds of the simple model, yet it makes up a small subspace compared to the model.

D. The refined model

The refined model was used by segmenting the set of 174 easy cases using the simple rule-based method. The RGB values for the resulting skin pixels from all the images were used to build a three dimensional histogram with one bin for each of the $256 \times 256 \times 256$ possible RGB values.

The histogram was then used to build the Boolean volumetric model shown in Fig. 1 by defining all RGB points as skin if the histogram count was above a limit arbitrarily set to 10. This step eliminated much of the noise.

Fig. 4 shows the resulting skin color model from different angles in HSV-space, RGB-space and CIElab space. The HSV renderings of the model show that the skin color pixels are relatively concentrated in the red part of the color wheel with medium saturation levels. The largest variation is along the brightness dimension. The shape of the volume is similar to that of a quarter ellipsoid with a tip toward the high side of the brightness dimension. The shape can be interpreted as follows. The brighter the skin, the less varied is the saturation and hue, while the darker the image pixels the larger the variation in hue and saturation is.

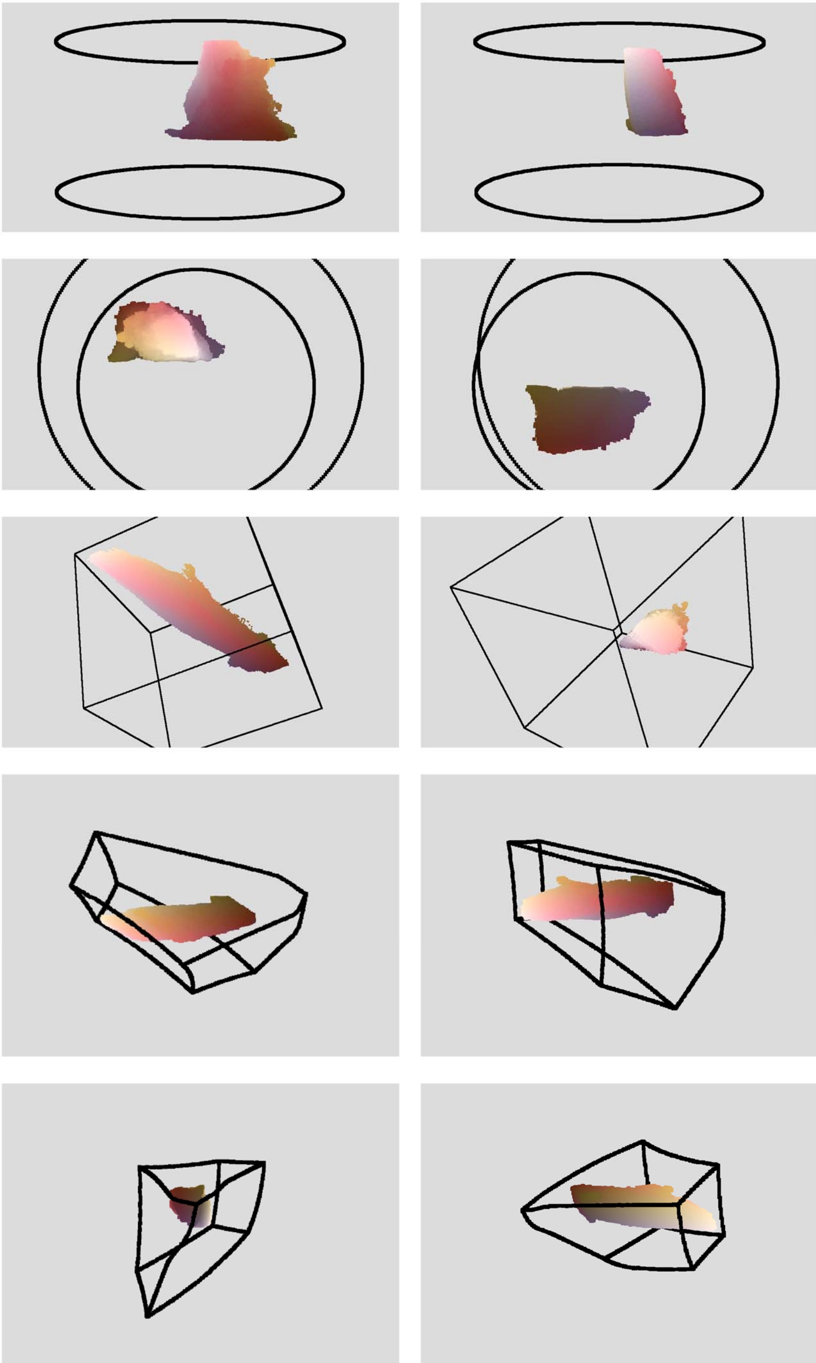

Figure 4. The refined skin detection model visualized. The four first images shows the model in HSV space viewed from two sides, top and bottom, the two next images shows the model viewed RGB space from the side and along the diagonal of the color cube. The last four images shows the model from various angle in the CIElab space.

The two RGB renderings of them model show a long, relatively thin triangularly shaped volume that is aligned parallel to the diagonal of the color cube going from the black corner to the with corner. The triangle can be interpreted as the hue and saturation variations of skin while the long shape indicates the intensity variations of skin. This is as expected, namely that the range of valid colors in terms of saturation and hue are limited but the main variation is in intensity. This observation supports the use of a static model.

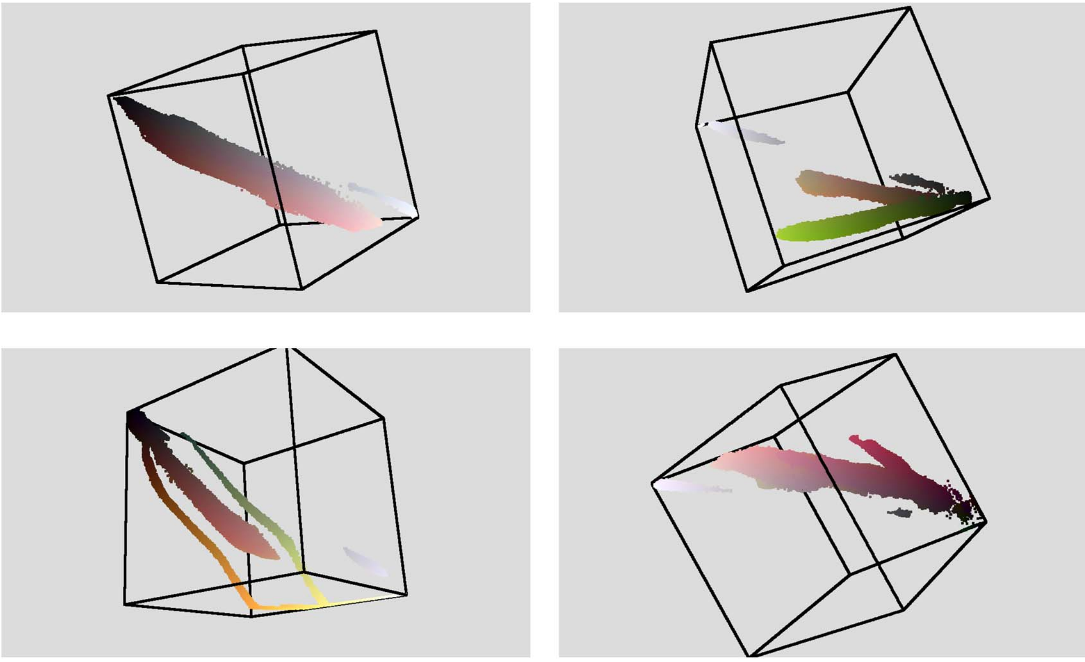

Figure 5. Volumetric cloud for all color pixels for a) easy case with black background, b) easy case with green background, c) difficult case with saturated lightbulb background and d) difficult case with hand colored background.

The last four images in Fig. 4 shows the resulting skin detection model in the perceptually uniform CIElab space. A visual inspection of the shape reveals that it is similar to the corresponding shape in RGB space. The shape of the skin color volume does not coincide with the regions of the CIElab space with the largest perceptually variations compared to the simple RGB model. The skin volume is somewhat in the middle of the CIE-lab space. Consequently, it appears that CIElab does not pose any obvious advantages over the other more simple color spaces. Its high computational complexity and unfamiliarity among programmers makes CIElab a less attractive..

III. Evaluation

Fig. 5 shows the RGB point clouds for all the pixels in four images representative of the four classes of images in our skin database. The top two visualizations of easy cases and the two bottom visualizations depict hard cases. All the images have a white identifying label in one of the corners and these labels causes the white “leaf” in the white corner of the color cubes. These leaves are outside the skin volume and are therefore effectively discarded.

The top left visualization in Fig. 5 is the simplest case, a hand on a close-to-black background. The background is successfully discarded as it is outside the bounds of the skin volume. The top right image shows visualization of the hands on a green background. The green colors follow a path in a different direction than the skin volume, and is successfully filtered since it is outside the volume. The visualization also shows a third “arm” of dark pixels belonging to the background that are also outside the bounds of the skin volume.

The bottom left visualization represents hands with a light bulb in the background. The resulting highly saturated yellow, orange and green colors follows clearly visible “bands” through the RGB color space. Since these bands are outside the bounds of the skin volume they are successfully filtered using the lookup table.

Finally, the bottom right visualization represents hands on a pink background that is similar to the color of the hands. The visualization shows a clearly visible dark red arm that falls outside the bounds of the skin volume and have therefore successfully filtered. There is also a small island of dark pixels towards the green side of the spectrum that are successfully filtered. However, some of the background pixels fall inside the bounds of the skin volume and will therefore trigger false positives. For this type of images, the pixel based information needs to be combined with spatial features in order to achieve a completely robust skin detection.

Fig. 6 shows skin detection for the two difficult cases in the above example. The left shows the original image (with identifiable information removed), the middle images shows the results after applying Peer’s rules and the right images shows the results using table-lookup.

The results show that the simple rule based method reports too many false positives shown by the orange band in the middle of the figure. The table lookup method successfully detects the skin area in this image. However, the image with the pink background is not successful with either of the methods. However, the number of false positives is lower with table-lookup compared to the simple rules.

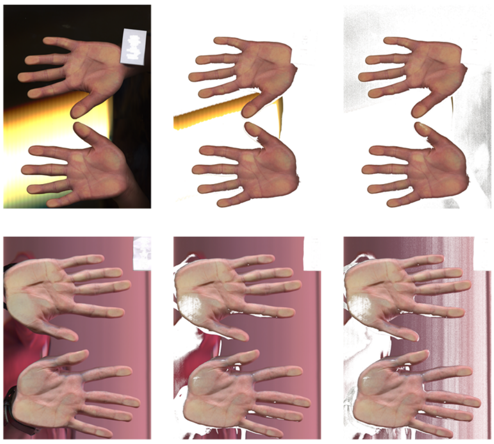

Figure 6. Skin detection with Peer's metod and table-lookup.

IV. Limitations

The skin detection model proposed herein assumes that the scenes captured in the images are lit by balanced, or close to white, light sources. Obviously, the perceived color of an object depends on a combination of the actual surface color and the color of the light shining on the object. The light shining on the object is a combination of direct light from light sources, ambient light and light reflected from other objects. If the light is monochrome, the resulting surface will appear with the same monochrome hue. The method thus, will not work in such environments, such as in nightclubs or concert venues with colored lights or images taken in streets at night light by sodium vapor bulbs that emit close to monochrome light. For such applications tailor made processing steps may need to be introduced as is often the case with image and video analysis systems [20, 21].

V. Conclusions

This study approached the classic skin detection problem from a visualization perspective. The results suggest that reliable pixel based skin detection can be performed simply, practically, efficiently and accurately using simple lookup tables in the RGB domain. The results confirm that skin is has a relatively limited variation in saturation and hue while the major variations occur in brightness. The results also suggest that a static model suffices for classifying pixels in many applications. Moreover, the results do not support the use of more complex color spaces that are similar to the human visual system. We therefore argue for performing skin detection in the hardware-centric RGB domain, as it is more familiar to programmers. The method is not capable of handling exceptional cases where skin is lit by non-balanced light, directly from monochrome light source or via reflections from other objects in the environment that only reflect light of certain frequencies.

References

[1] M. M. Hasan, J. M. Thakur, and P. Podder, "Face Detection in Color Images Using Skin Color." International Journal of Scientific & Engineering Research, Vol. 5, No. 6, 2014.

[2] S. A.-K. Hussain, "Human Face Detection with skin color properties.", International Journal of Scientific & Engineering Research, Vol. 6, No. 6, 2015.

[3] M. K. Ahuja, and A. Singh, "Principal Component Analysis based Hand Gesture Recognition" International Journal of Advanced Research in Computer Engineering & Technology, Vol. 4, no. 7, 2015.

[4] F. E. Sandnes, "Measuring 2D:4D finger length ratios with Smartphone Cameras", In: Proceedings of IEEE SMC 2014, IEEE Press, pp. 1712-1716, 2014.

[5] F. E. Sandnes, "An Automatic Two-hand 2D:4D Finger-ratio Measurement Algorithm for Flatbed Scanned Images," In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE Computer Society Press, pp. 1203-1208, 2015.

[6] F. E. Sandnes, "A Two-Stage Binarizing Algorithm for Automatic 2D:4D Finger Ratio Measurement of Hands with Non-Separated Fingers", In: Proceedings of 11th International Conference on Innovations in Information Technology (IIT'15), IEEE, pp. 178-183, 2015.

[7] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques" In Proceedings of Graphicon, Vol. 3, 85-92, 2003.

[8] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods" Pattern recognition, Vol. 40, No. 3: 1106-1122, 2007.

[9] J. Kovač, P. Peer, and F. Solina, "Human skin color clustering for face detection", In: Proceedings of 2003 International Conference on Computer as a Tool, IEEE, 2003.

[10] A. Nusirwan, C. W. Kit, and S. John, "RGB-H-CbCr Skin Colour Model for Human Face Detection", In: Proceedings of The MMU International Symposium on Information & Communications Technologies, pp. 579-584. 2006.

[11] R. Khan, Z. Khan, M. Aamir, and S. Q. Sattar, "Static filtered skin detection", International Journal of Computer Science Issues, Vol. 9, No. 2, 257-261, 2012.

[12] A.-T., Z. Hamid, R. Wirza, O.K. Rahmat, M. I. Saripan, and P. S. Sulaiman, "Skin Segmentation Using YUV and RGB Color Spaces", JIPS, Vol. 10, No. 2, 283-299, 2014.

[13] G. Gomez, and E. Morales, "Automatic feature construction and a simple rule induction algorithm for skin detection", In: Proceedings of the ICML workshop on Machine Learning in Computer Vision, pp. 31-38. 2002.

[14] N. A. Ibraheem, and R. Z. Khan, "Multiple Histogram Technique for Robust Skin Color Based Segmentation", American Journal of Engineering Research, Vol. 2, No. 5, 2013.

[15] A. Nadian-Ghomsheh, "Color Constancy For Improving Skin Detection", International Journal of Image Processing, Vol. 8, No. 6, 479, 2014.

[16] T. M. Mahmoud, "A new fast skin color detection technique." World Academy of Science, Engineering and Technology, Vol. 43, 501-505, 2008.

[17] B. D. Zarit, B. J. Super, and F. K. H. Quek, "Comparison of five color models in skin pixel classification", In: Proceedings of International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, pp. 58-63. IEEE, 1999.

[18] Q. Chen, H. Wu, and M. Yachida, "Face detection by fuzzy pattern matching." In: Proceedings of The Fifth International Conference on Computer Vision, pp. 591-596. IEEE, 1995.

[19] D.-H. Xu, A. S. Kurani, J. D. Furst, and D. S. Raicu. "Run-length encoding for volumetric texture", Heart, Vol. 27, 25, 2004.

[20] Y.-P. Huang, T.-W. Chang, Y.-R. Chen, and F. E. Sandnes, "A back propagation based real-time license plate recognition system", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 22, No. 2, 233-251, 2008.

[21] Y.-P. Huang, L.-W. Hsu, and F.-E. Sandnes, "An intelligent subtitle detection model for locating television commercials", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 37, No. 2, 485-492, 2007.