RESEARCH ARTICLE   OPEN ACCESS

Region based elimination of noise pixels towards optimized classifier models for skin pixel detection

Gagandeep Kaur*, Seema Pahwa**, Kushagra Sharma***

*(Department of IT, Sri Sukhmani College of Engineering, & Technology, Derabassi- Punjab Technical University)
** (Department of IT, Sri Sukhmani College of Engineering, & Technology, Derabassi - Punjab Technical University)
*** (Department of Computer Science and Engineering, Sri Sukhmani College of Engineering, & Technology, Derabassi - Punjab Technical University)

ABSTRACT
The extraction of the skin pixels in a human image and rejection of non-skin pixels is called the skin segmentation. Skin pixel detection is the process of extracting the skin pixels in a human image which is typically used as a pre-processing step to extract the face regions from human image. In past, there are several computer vision approaches and techniques have been developed for skin pixel detection. In the process of skin detection, given pixels are been transformed into an appropriate color space such as RGB, HSV etc. And then skin classifier model have been applied to label the pixel into skin or non-skin regions.
Here in this research a "Region based elimination of noise pixels and performance analysis of classifier models for skin pixel detection applied on human images" would be performed which involve the process of image representation in color models, elimination of non-skin pixels in the image, and then pre-processing and cleansing of the collected data, feature selection of the human image and then building the model for classifier. In this research and implementation of skin pixels classifier models are proposed with their comparative performance analysis. The definition of the feature vector is simply the selection of skin pixels from the human image or stack of human images. The performance is evaluated by comparing and analysing skin colour segmentation algorithms. During the course of research implementation, efforts are iterative which help in selection of optimized skin classifier based on the machine learning algorithms and their performance analysis.

Keywords-Digital Image Processing, Data Mining, Digital Image Processing, Skin Pixel Detection, Color Models, WEKA

I.   INTRODUCTION

The identification and determination of the face in a human image is called face detection. Face detection is a research problem to select the sub-window of human image which contains a face. In computer vision, there are various applications such as face processing. Several applications, such as face processing, computer human interaction, human crowd surveillance, biometric, video surveillance, artificial intelligence and content-based image retrieval [1]. Face detection is a pre-processing step in all these application which is used to obtain the "object" related to the face region. In the form of face detection, the proposed processes in research can be defined as shown in the following figure 1.

In research, there are many techniques are proposed for pre-identification of the face in a human image which start from the phase of skin color segmentation. The skin color segmentation is representing the human image in a color models. In computer vision, there are many color models to represent the image; these are RGB, YCbCr, HSV,

and LAB color models. Next step is to convert the skin image into binary image which contains the skin regions. Then next step in face detection is to remove non human face area from the segmented image with the help of the Knowledge-based methods or human face features [1]. Here in this research, the elimination of non-skin region from the skin image and selection of skin part with the help of representation of human in particular color model such as RGB and HSV is presented. Then machine learning algorithms are applied to build an optimized skin classifier on the segmented image.

Skin Segmentation based on Color Models:

To represent the human image, color model is a useful piece of information which decide the optimality and efficiency of the skin segmentation. For skin detection and segmentation, the color model is important and useful parameter [2]. Like, in terms of face detection, based on the appropriate color model representation, the extensive search of human

face pixels can be avoided which help in deciding the human image region related to face.
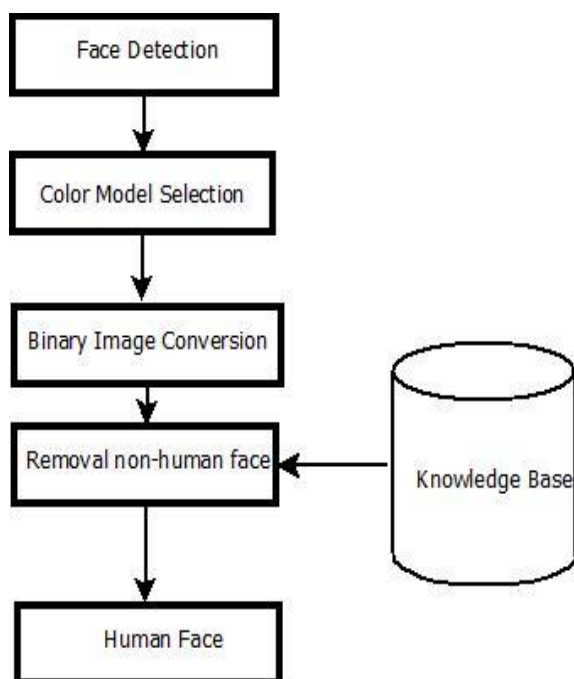

Figure 1: Face Detection approach

In this step, we describe that how non skin color is rejected from an Image so that the image may contains only skin like areas, which will be our skin color segmented image for further processing. From different type of color models, in HSV color model, Hue (H) is not reliable for the discrimination task when the saturation is low, Also in YCbCr color model, the distribution of skin areas is consistent across different races in the Cb and Cr color spaces, the RGB color model is lighting sensitive so Therefore, when we use different color models under uncontrolled conditions, and we get consequently result for skin color detection. The accuracy of skin detection depends on both the color model and the method of skin pixels classification or detection. Hence, the challenge problem is how to select color models that are suitable for skin pixel classifications under different varying conditions. In this thesis, there are four color models are used for skin color segmentation or detection of skin pixels. These are RGB, YCbCr, and HSV and CIELAB color models. The combination of these color models overcomes all varying lighting conditions and changes in illumination, and it gives better result than individual color model result [3]. Figure 2 shows the combination of skin color segmentation.

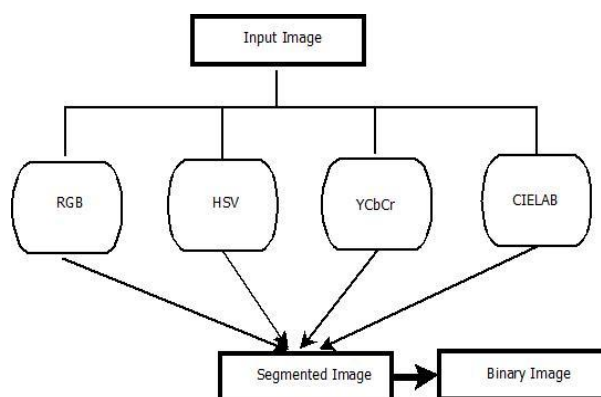RGB YCbCr HSV CEIL*a*b* Final Skin Color Segmented Image Binary Image Input image

Fig 2: Skin color segmentation

3. BINARY IMAGE FORMATION

As stated above, the segmented skin image is only contained the skin areas or skin regions but there is no information which region or which area is having human face or part of human skin, thereby we need to send this image for further classification stage which reject the non-human face of the image and select the part of the image which match with the human face. The similar principle is working on the skin detection which is majorly divided into two broad categories as skin region selection which is basically the elimination of non-skin regions and other is building the classification model to select only skin related pixels.

II. Skin Classifier Model
A. Scope and Objectives

The research and development involved in this research include:

➢ Literature Survey related to skin detection.
➢ Proof-of-Concept of Skin Pixel Detection
➢ Elimination of non-skin based on region based selection
➢ Data pre-processing and feature selection of human image.
   o Feature selection.
   o Classification
   o Model Creation
➢ Development of optimized skin classifier based on multiple iterations.
➢ Experimental Results Illustrations and future scope.
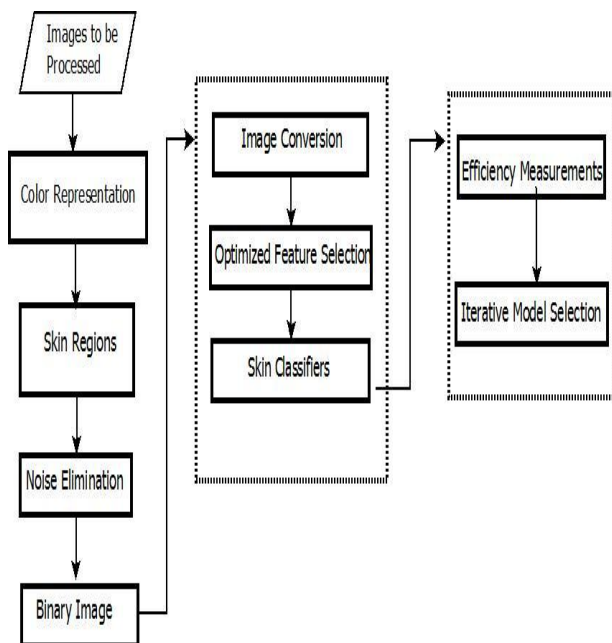
Methodology Used:

Figure 3: Flowchart and Algorithm

> ➢ **Tools and Techniques:**
>   o Weka:
>   o ImageJ aka Fiji

ImageJ aka Fiji:

ImajeJ: Fiji is an image processing package. It can be described as a distribution of ImageJ (and ImageJ2 together with Java, Java3D and a lot of plugins organized into a coherent menu structure. Fiji compares to ImageJ as Ubuntu compares to Linux.

Weka supports several standard data mining tasks, more specifically, data pre-processing, clustering, classification, regression, visualization, and feature selection.

The main goal of this plugin is to work as a bridge between the Machine Learning and the Image processing fields. It provides the framework to use and, more important, compare any available classifier to perform image segmentation based on pixel classification.

1. Data Collection: First of all the data will be collected related to human images and segregated from different sources.
2. Data Filtering and Cleansing: The data we have collected are not clean and may contain errors, missing values, noisy or inconsistent data. So we need to apply different techniques to get rid of such anomalies.
3. Feature Selection: Features will be selected based on human images which related to pixels of the images.
4. Training the model/model creation: The model will be created and trained based on the training data-set.

5. Testing the model: Testing of the model based on supplied test data
6. Evaluation: Performance evaluation of the developed model.

Modules of Implementation:

1. Pre-processing: In this step, the data from various sources related to human image will be collected and pre-process to fuse it in the format of Weka tool which is .arff format. When a color image is given as an input to the algorithm, the first step is pre-processing. This step is mainly concerned with the image conversion and cropping.
2. Feature Extraction: The data set collected will further be processed to select the relevant features of the human image. Feature extraction transforms the input data into the set of features Feature extraction is concerned with recovering the defining attributes obscured by imperfect measurements.
3. Classification: For classification, we train the machine with the help of learning algorithms.
4. Model Creation: Finally the accurate model will be created to test and cross-validate the results.
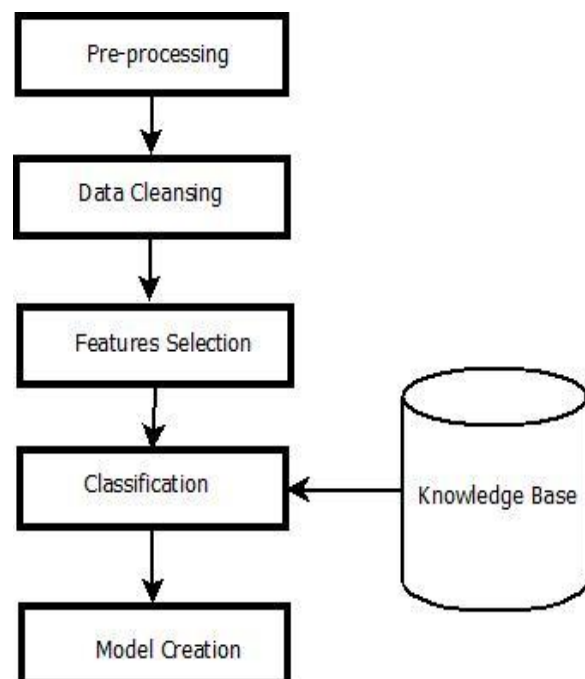

Figure 2: Modules of implementation

Machine learning algorithms implemented and compared to build optimized classifier model:

1. Naive Bayes Classifier: It is a simple probabilistic model of machine learning which is used to develop a classifier.
2. Decision Tree: It is basically a predictive model of machine learning subject.

3. SMO: Sequential minimal optimization (SMO) is an algorithm for solving the optimization problem.
4. J48: C4.5 is an algorithm used to generate a decision tree developed by Ross Quinlan. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier.
5. ZeroR: ZeroR is the simplest classification method which relies on the target and ignores all predictors. ZeroR classifier simply predicts the majority category (class). Although there is no predictability power in ZeroR, it is useful for determining a baseline performance as a benchmark for other classification methods.

Experimental Results:

Algorithm	TP	FP	Precision	Recall	ROC area
Decision Tree	0.949	0.205	0.948	0.949	0.917
Bayesian Logistic Regression	0.938	0.251	0.936	0.938	0.843
J48	0.958	0.187	0.957	0.958	0.949

Table 1: Comparison of classifier models

Classification Rates: (Total instances=2067)

Algorithm	Correctly classified instances	Incorrectly classified instances
Decision Tree	1962 (94.9202%)	105(5.0798%)
Bayesian Logistic Regression	1938(93.7591%)	129(6.2409%)
J48	1980(95.791%)	87 (4.209%)

Table 2: Classifications rates of classifiers

Confusion Matrix (Decision Tree):
```
a           b         ←-------classified as
1748        47        a=class 1
58          194       b= class 2
```

Confusion Matrix (Bayesian Logistic Regression):
```
a           b         ←-------classified as
1757        58        a=class 1
71          181       b= class 2
```
Confusion Matrix (J48):
```
a           b         ←-------classified as
1781        34        a=class 1
```
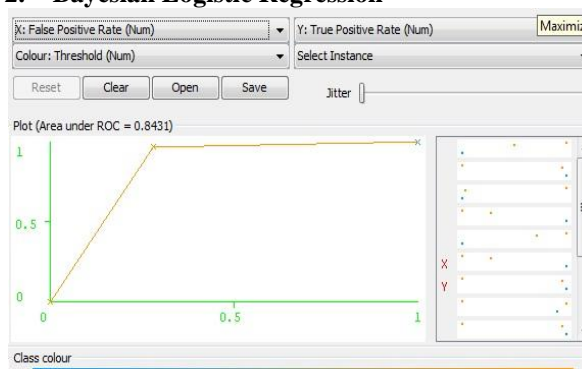
53   199   b= class 2

ROC curves:
1. Decision Tree
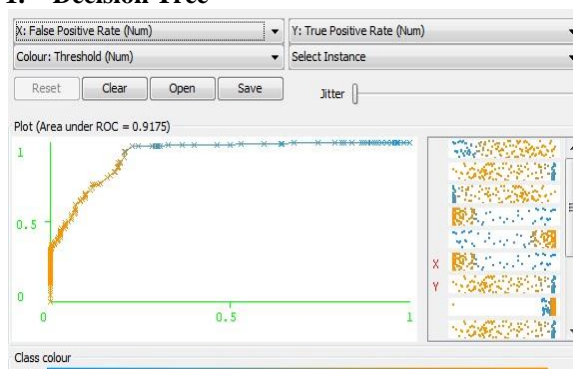
Figure 3: ROC curve of Decision Tree

2. Bayesian Logistic Regression

Figure 4: ROC curve of Bayesian Logistic Regression
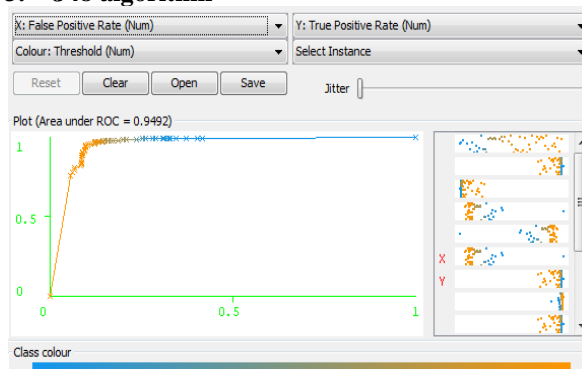
3. J48 algorithm

Figure 5: ROC curve of J48

III. CONCLUSION

Here in this research, the machine learning based approach to build the skin classifier is being presented. The iterative mechanism is implemented to build the optimized skin classifiers. First step in the research is elimination of noise pixels and selection of skin-regions or skin areas which is

further send to the second phase to process it further. The features which are taken for skin classifier need to be optimized in the future.

IV. ACKNOWLEDGEMENTS

References

[1] Ming-Hsuan Yang, member, IEEE, David J. Kriegman, Senior Member, IEEE, and Narendra Ahuja, Fellow, IEEE. "*Detecting Faces in Images: A Survey*", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. no. 1, January 2002.

[2] Ah J. Liu and Y. H. Yang, "*Multiresolution Color Image Segmentation*", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 7, pp. 689-700, 1994.

[3] Sanjay Kr. Singh1, D. S. Chauhan2, Aayank Vatsa, Richa Singh "*A Robust Skin Color Based Face Detection Algorithm*", Tamkang Journal of Science and Engineering, Vol. 6, No. 4, pp. 227-234, 2003.

[4]. Devendra Singh Raghuvanshi et al. Human Face Detection by using Skin Color Segmentation, Face Features and Regions Properties, International Journal of Computer Applications (0975 – 8887) Volume 38– No.9, January 2012

[5] Fu You-jia et. al, Rotation Invariant Multi-View Color Face Detection Based on Skin Color and Adaboost Algorithm, 2010

[6] Wang Jianguo, Wang Jiangtao, Yang Jingyu, "*Rotation-invariant face detection in color images with complex background*," Computer Engineering, Shanghai, vol. 34, pp. 210-212, February 2008. (in Chinese)

[7] Zhang Hongming, Zhao Debin, Gaowen, "*Face detection under rotation in image plane using skin color model, neural network and feature-based face model*," Chinese Journal of Computers, Beijing, vol. 25, pp. 1250-1256, November 2002. (in Chinese)

[8] Paul Viola, Michael J. Jones, "*Rapid object detection using a boosted cascade of simple features*," In: Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, USA. 2001.

[9] Stan Z. Li, Zhang Zhenqiu, "*FloatBoost learning and statistical face detection*," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 26, pp. 1112-1123, September 2004.

[10] Huangchang, Ai Haizhou, Liyuan, Lao Shihong, "*Vector boosting for rotation invariant multi-View face detection*," In Proceedings of the Tenth IEEE International [11] Conference on Computer Vision, Beijing, China, Oct. 2005. Zhou Minquan, Geng Guohua, Weina, The Technology of Content-based Image Retrieval, Beijing: Tsinghua University Press, 2007, pp. 22-23. (in Chinese)

[12] Zhang Yujin, Image Engineering, vol. I. Beijing: Tsinghua University Press, 1999, pp. 68-70. (in Chinese)

[13] Stan Z. Li, Zhulong, Zhang Zhenqiu, Zhang Hongjian, "*Learning to detect multi-view faces in real-time*," In Proceedings of the 2nd International Conference on Development and Learning, Washington DC, June, 2002.

[14] *Gagandeep Kaur et al,* Data Mining Based Skin Pixel Detection Applied On Human Images: A Study Paper *Int. Journal of Engineering Research and Applications, ISSN : 2248-9622, Vol. 4, Issue 7(Version 5), July 2014, pp.48-53.*

[15] Bernhard Fink, K.G., Matts, P.J.: *Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces.* Evolution and Human Behavior 27(6) (2006) 433–442

[16] Ahmed Elgammal, *Crystal Muang and Dunxu Hu: Skin Detection - a Short Tutorial.*

[17] Wei Ren Tan, Chee Seng Chan, Pratheepan Yogarajah, and Joan Condell (2012), "*A Fusion Approach for Efficient Human Skin Detection*", IEEE Transactions on Industrial Informatics, Vol.8, No.1,pp.138-147.

[18] Chang-Yul Kim, Oh-Jin Kwon, and Seokrim Choi (2011), "*A Practical System for Detecting Obscene Videos*",IEEE Transactions on Consumer Electronics, Vol.57, No.2, pp. 646-650.

[19] Fumihito Yasuma, Tomoo Mitsunaga, Daisuke Iso, and Shree K. Nayar (2010)," *Generalized Assorted Pixel Camera: Postcapture Control of Resolution, Dynamic Range and Spectrum*", IEEE Transactions on Image Processing, Vol.19, No.9,pp.2241-2253.

[20] Jalal A.Nasiri, H. Sadoghi Yazdi and Mohmoud Naghibzadeh (2008), "*A Mouth*

Detection Approach Based on PSO Rue Mining on Color Images", 5th Iranian Conference on Machine Vision and Image Processing, November 4-6.

[21] Nils Janssen and Neil Robertson (2008), "*On the detection of low-resolution skin regions insurveillance images*", The Eighth International Workshop on Visual Surveillance.

[22] Abbas Cheddad, Joan Condell, Kevin Curran and Paul Mc Kevitt (2009), "*A New Color space for skin tone detection*", ICIP, pp.497-500.

[23]. SKIN SEGMENTATION USING COLOR AND EDGE INFORMATION *Son Lam Phung, AbdesselamBouzerdoum, and Douglas Chai,* School of Engineering and Mathematics, Edith Cowan University Perth, Western Australia, AUSTRALIA, 2012

[24]. J. Liu and Y. H. Yang, "*Multiresolution Color Image Segmentation*", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 7, pp. 689-700, 1994.

[25]. Devendra Singh Raghuvanshi and DheerajAgrawal, *Human Face Detection by using Skin Color Segmentation*, Face Features and Regions Properties, *International Journal of Computer Applications (0975 – 8887), Volume 38–No.9, January 201.*

[26]. M.J. Jones and J.M. Rehg, "*Statistical Color Models with Application to SkinDetection,*" Int'l J. Computer Vision, vol. 46, no. 1, pp. 81-96, Jan. 2002.

[27]. J. Brand and J. Mason, "A *Comparative Assessment of Three Approaches toPixel-Level Human Skin Detection,*" Proc. IEEE Int'l Conf. Pattern Recognition,vol. 1, pp. 1056-1059, Sept. 2000.