Facial Feature Extraction for Face Modeling Program

Ha-sung Koo(hskoo@hanseo.ac.kr), Ho-guen Song(hksong@hanseo.ac.kr)

Abstract: - In this paper suggests to define 20 facial features being used in 3D face modeling program and the method of extraction of 20 facial features from 2D image. The experimental image is to be restricted to input one person's image, background of the image can be taken in anywhere, the lighting condition is not uniform, and there is no racial restriction. The suggested method is to seek facial candidate region by Harr Classifier and to decide eye candidate region and extract eye features by dilate operation then decide lip candidate region using the features. The relative color difference of a* in the L*a*b* color space was used to extract lip feature and to seek nose candidate region and detected 20 features from 2D image by analyzing end of nose.

Keywords— facial features extraction, L*a*b* color space, Harr classifier, no racial restriction.

I. INTRODUCTION

Many practical studies to extract feature from face such as Video Indexing, converting 2D face to 3D face, facial predictions are under proceeding.

It is important for such technology how to extract precise futures from face in initial stage. Even if same algorithm, different result will be taken depending on background of the image, lighting condition and number of people. Therefore it is necessary to develop an algorithm being suitable to practical field than the algorithm which satisfies all occasions.

This study is to define 20 features as a first stage to convert 2D face to 3D face to find lip and nose exactly regardless race and lightings.

The existing studies of finding the position of eye and lip are the method to extract facial feature and eye feature using color and to recognize eye using evaluate function reflecting symmetry of two eyes, method[1] to find the position of eye using mask reflecting gray-level distribution between the apple of the eye and the white of the eye. method[2] to find a pair of eyes until fulfill fixed and restricted conditions on the facial region by changing threshold and repeating iteration of bi-level thresholding, method[3]to find eye and mouth using eigen face and eigen mouth which extended the conception of eigen face. These studies lay emphasis to the process of recognition of eye and recognize rough position of mouth, method [4] as a study to seek lip shape and to seek the edge from facial region, and to recognize lip contour line by checking corner angle, direction, and curvature. Method [5] to extract lip region using color information and to approximate the exterior angle line of upper and lower lip as biquadratic and quadratic. Method to recognize lip using Kalman filter and lip color model, to find lip region using color information and HMM(Hidden MarKov Model) and Method[7] to recognize exterior line lip using Snake method.

The experiment image of this thesis are own facial photograph taken by generally used mobile phone camera and the facial photograph obtained by internet for getting various racial photograph. Only one person's face is exist in the acquisition image and there is no restriction in the background, lighting and resolution of photograph. The object of the embodied purpose of developing the system is own face photograph taken by mo

This Study used the classifier which extracts face candidate region from the given inputted images, find eye candidate region using Dilate Operation and after extracting the end point of eye thru regional Binary among the candidate region been found, extracted lip among the lip candidate regions appeared by the connection of two eyes. The extraction of lip among lip candidate region could extract racial strong lip features using the relative color difference in the L*a*b* color space which has not been suggested up to now. After extracting eye and lip, to seek nose candidate region and analyzed 20 features out from 2D facial images by analyzing end of nose. Figure 1 is the diagram of the algorithm being suggested in this Study. The figure 2 is 20 facial features seeking from 2D facial image. These 20 features are 8 features from "a" to "g" on eye, 3 features from "i" to "k" on nose, 4 feature from "l" to "o" on lip, five feature from "p" to "t' on outline of face. These features are defined features as the primary study stage to convert into 3D face matching to the face shape mash being manufactured.

II. EXTRACTING FACE REGION AND FEATURE OF EYES

II.1 EXTRACTING FACE REGION

In this paper we used face detector that developed by Paul-Viola and Micheal-Jones, to extract face region from experiment images. The Viola-Jones' s detector do AdaBoot in each nodes of cascades to get high detecting rates with low rejecting ratio. The algorithm do single decision as Method(1), and also do classified function as Method(2). sign() function returns negative numbers as -1, positive numbers as +1 and 0 as 0. This algorithm used reference paper[9]. Viola - Jones carry out AdaBoost in each node of cascades for getting high rate of detection with low rejection rate.

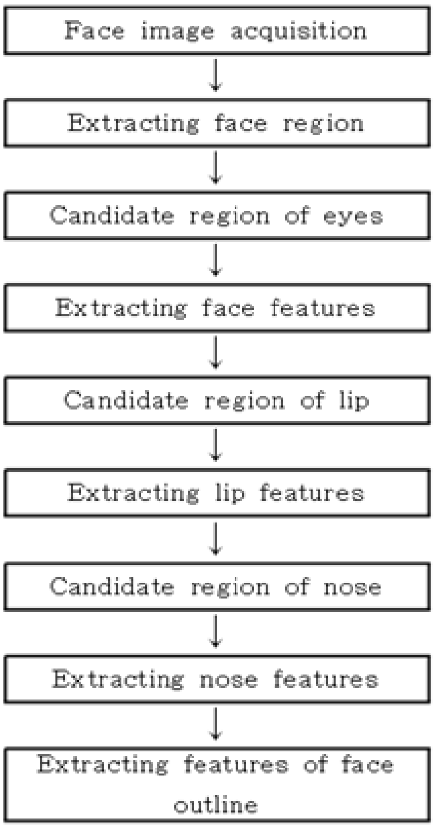


Fig. 1. Block diagram of proposed algorithm

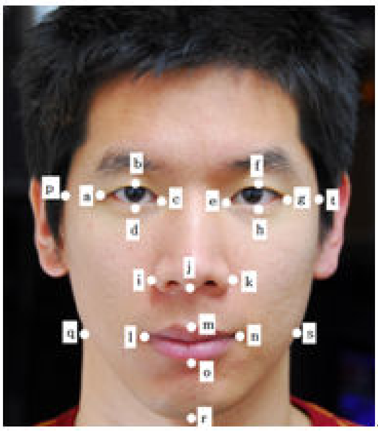


Fig. 2. 20 features used in this paper

$$fi = \begin{cases} +1 & vi \geq ti \\ -1 & vi < ti \end{cases} \tag{1}$$

$$F = \text{sign}(w1f1 + w2f2 + \dots wnfn) \tag{2}$$

The figure 3 is example of extracting candidate face region.

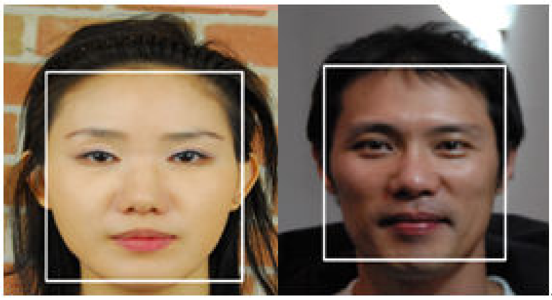


Fig. 3. Example of extracting candidate face region

II.2 EXTRACTING FEATURE OF EYES

Set up the upper left side as left eye candidate region by dividing face region by four which was found from figure 2 and figure 3 shows eye candidate region

The eyes are generally known as the easiest region to extracting face characteristics because eyes has the highest brightness differences. However if do binary it with brightness, it also binary the eyebrows too and brings statistics from hair so that make things hard to find the correct eye region. Therefore to make decision eye region with deviation value of eyebrow which is small by labeling after bi-level thresholding, and there is a method of extraction using evaluate function in hybrid with different from the method which makes decision with the number of edge by Canny edge.

The Canny edge detection algorithm is known to many as the optimal edge detector. Canny's intentions were to enhance the many edge detectors already out at the time he started his work. He was very successful in achieving his goal and his ideas and methods can be found in his paper, "*A Computational Approach to Edge Detection*". In his paper, he followed a list of criteria to improve current methods of edge detection. The first and most obvious is low error rate. It is important that edges occurring in images should not be missed and that there be NO responses to non-edges. The second criterion is that the edge points be well localized. In other words, the distance between the edge pixels as found by the detector and the actual edge is to be at a minimum. A third criterion is to have only one response to a single edge. This was implemented because the first 2 were not substantial enough to completely eliminate the possibility of multiple responses to an edge.

But in case of Europeans having deep-set eyes, these kinds of method has demerits which is difficult to extract from 2D image which looks eye and eyebrow are close to each other. Expression (1) means dilation, in this expression, the part of which the comparison of light and shade is big is expended and has more value than original value, so if divide by original value, can get big value at the meeting space of dark and bright which is suitable to extract eyeball. This is based on the assumption that at least the comparison of light and shade of eyeball is the biggest part in facial candidate region. Expression (3) is EyeMapL. Isolated eyeball can be extracted by calculating EyeMapL from eye candidate region with the extraction of 10% of high level after generalization from 0 to 255. By calculating the width of extracted eyeball and with it as center and setting up the width region with 5 times of the height of the eyeball and extract the end point of eye thru local thresholding. Figure 4 below is candidate region of eyes and figure 5 is the process of extracting eyeball.

$$EyeMapL = \frac{Y(x,y) \oplus g_{\sigma}(x,y)}{y(x,y)} \quad (3)$$

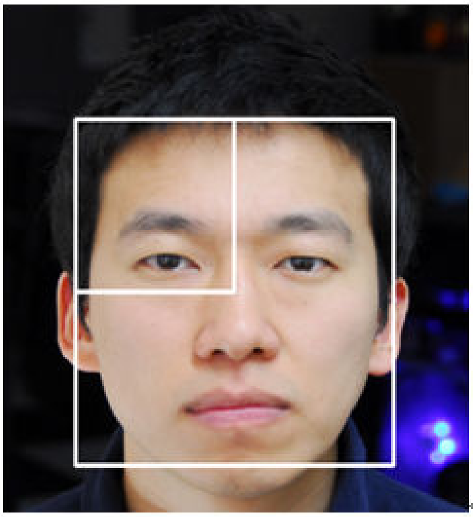
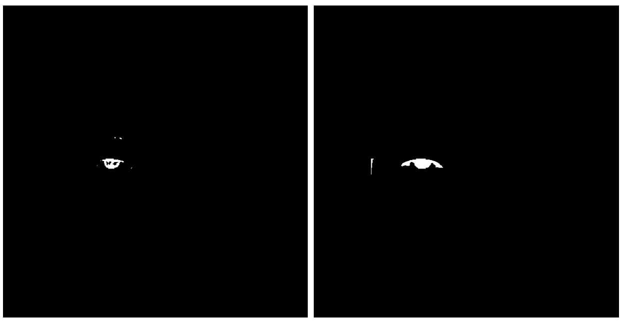


Fig. 4.Candidate region of eyes



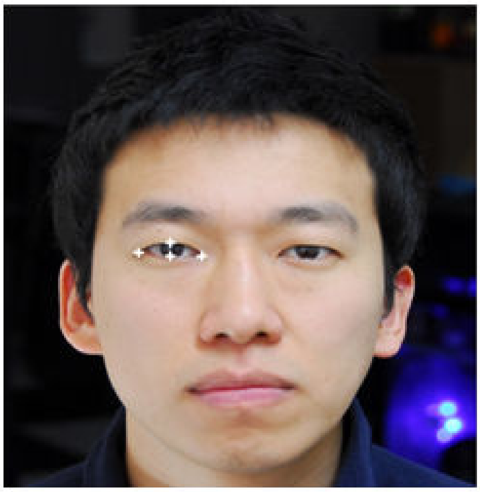


Fig. 5. Process of extracting feature of eye

III. Extracting features lip and nose

III.1. *COLOR SPACE*

III.1.1. YcbCr Color Space

YCbCr or Y′CbCr, sometimes written YCbCr or Y′CbCr, is a family of color spaces used as a part of the color image pipeline in video and digital photography systems. Y′ is the luminance component and Cb and Cr are the blue-difference and red-difference chroma components. Y′ (with prime) is distinguished from Y which is luminance, meaning that light intensity is non-linearly encoded using gamma correction.

Y′CbCr is not an absolute color space, it is a *way of encoding* RGB information. The actual color displayed depends on the actual RGB colorants used to display the signal. Therefore a value expressed as Y′CbCr is only predictable if standard RGB colorants or an ICC profile are used. Cathode ray tube displays are driven by red, green, and blue voltage signals, but these RGB signals are not efficient as a representation for storage and transmission, since they have a lot of *mutual* redundancy.

YCbCr and Y′CbCr are a practical approximation to color processing and perceptual uniformity, where the primary colors corresponding roughly to Red, Green and Blue are processed into perceptually meaningful information. By doing this, subsequent image/video processing, transmission and storage can do operations and introduce errors in perceptually meaningful ways. Y′CbCr is used to separate out a luma signal (Y′) that can be stored with high resolution or transmitted at high bandwidth, and two chroma components (C_B and C_R) that can be bandwidth-reduced, subsampled, compressed, or otherwise treated separately for improved system efficiency.

One practical example would be decreasing the bandwidth or resolution allocated to "color" compared to "black and white", since humans are more sensitive to the black-and-white information .Figure 6 is YCbCr plane wit y=0.5.

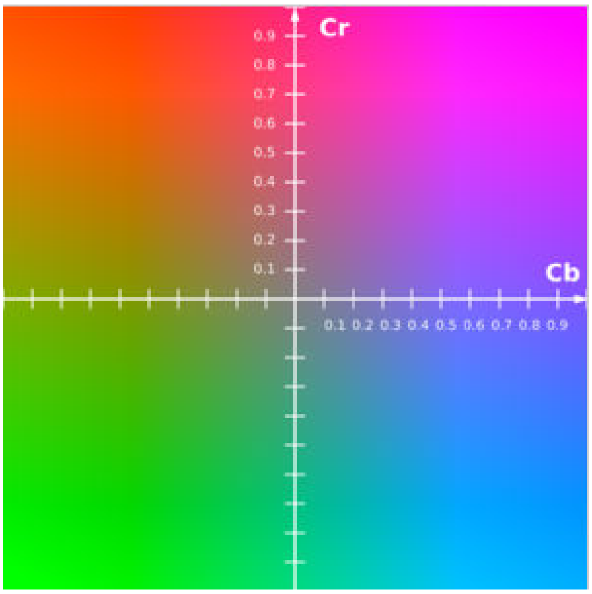


Fig. 6. YCbCr plane wit y=0.5

III.1.2. *L*a*b* Color Space*

CIE *L*a*b** (CIELAB**)** is the most complete color space specified by the International Commission on Illumination (*Commission Internationale d'Eclairage*, hence its *CIE* initialism). It describes all the colors visible to the human eye and was created to serve as a device independent model to be used as a reference. The three coordinates of CIELAB represent the lightness of the color (***L**** = 0 yields black and ***L**** = 100 indicates diffuse white; specular white may be higher), its position between red/magenta and green (***a****, negative values indicate green while positive values indicate magenta) and its position between yellow and blue (***b****, negative values indicate blue and positive values indicate yellow). The asterisk (*) after *L*, *a* and *b* are part of the full name, since they represent L*, a* and b*, to distinguish them from Hunter's L, a and b, described below. Since the *L*a*b** model is a three-dimensional model, it can only be represented properly in a three-dimensional space. Two-dimensional depictions are chromaticity diagrams: sections of the color solid with a fixed lightness. It is crucial to realize that the visual representations of the full gamut of colors in this model are never accurate; they are there just to help in understanding the concept. Because the red/green and yellow/blue opponent channels are computed as differences of lightness transformations of (putative) cone responses, CIELAB is a chromatic value color space. Figure 7 is color space of L*a*b*. Unlike the RGB and CMYK color models, *Lab* color is designed to approximate human vision. It aspires to perceptual uniformity, and its *L* component closely matches human perception of lightness. It can thus be used to make accurate color balance corrections by modifying output curves in the *a* and *b* components, or to adjust the lightness contrast using the *L* component. In RGB or CMYK spaces, which model the output of physical devices rather than human visual perception, these transformations can only be done with the help of appropriate blend modes in the editing application.

Because *Lab* space is much larger than the gamut of computer displays, printers, or even human vision, a bitmap image represented as Lab requires more data per pixel to obtain the same precision as an RGB or CMYK bitmap. In the 1990s, when computer hardware and software was mostly limited to storing and manipulating 8 bit/channel bitmaps, converting an RGB image to Lab and back was a lousy operation. With 16 bit/channel support now common, this is no longer such a problem.

Additionally, many of the “colors” within Lab space fall outside the gamut of human vision, and are therefore purely imaginary; these “colors” cannot be reproduced in the physical world. Though color management software, such as that built in to image editing applications, will pick the closest in-gamut approximation, changing lightness, colorfulness, and sometimes hue in the process, author Dan Margulis claims that this access to imaginary colors is useful going between several steps in the manipulation of a picture

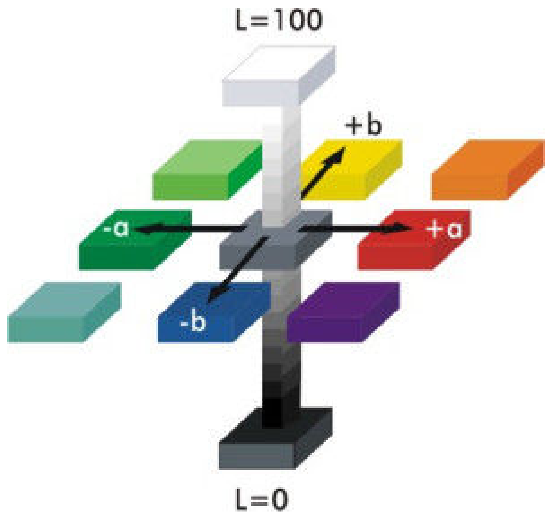


Fig. 7. Color space of La*b*.

III.2. EXTRACTING FEATURES OF LIP

After extracting eyes, to draw straight line between eyebrows and call the interval L. To draw regular triangle as much interval as L. Then call the other space which is not apex of eyebrow as center of lip candidate region. With it as center, the quadrilateral (2/3LX x 2/3LY) is to be lip candidate region. Suggesting Color Space for seeking lip candidate region. In case of lip, it has more red color than average in face so must use color information. But the degree of strength or weakness of redness is different in Blackman, Whiteman, man and woman and it has much difference depending on lighting condition and resolution of photograph. Current thesis use mainly HSI Color Space and YCbCr. HSI found the region of which Cr is higher than skin color using LipMap. In the space of HSI to judge lip region with the difference of H value. In the image of closed lip, the gray level of the meeting zone of the two lips is the most low, so with the Histogram find lip region too. But different results were carried out depending on lighting, resolution of photograph and race after studying various methods,. This thesis suggest L*a*b Color Space for finding lip region.

The extraction of lip region being suggested in this thesis pays attention to a* channel which can express the size between green color and red color at a time. After getting the average of a* on lip candidate region, to figure out deviation value. And the space which is bigger than 1/2 of deviation value is to be lip candidate. Figure 6 is candidate region of lip and the figure 6 shows lip candidate region and cumulative distribution of X axis in lip region using a*.

In the lip candidate region like figure 8, the biggest value in the Histogram of X- axis using a* is to be lip center region and the labeling region in the lip center region is to be the value of thickness of 2nd lip candidate region. The reason to find the 2nd lip candidate region is that the features which this study is trying to find include the points of lip-center and lip-end same as figure 1. In case of lip-end, there are many cases which is not included in the lip candidate of figure 4 and when close the lip, the color of lip-end appeared to be darker than other parts. Therefore 2nd ca야date region is determined like as figure 9 and the thickness of lip as previously obtained and the length of horizontal is fixed from "a" to "g" of figure 1. The black color of lip-end is obtained thru local thresholding in gray level

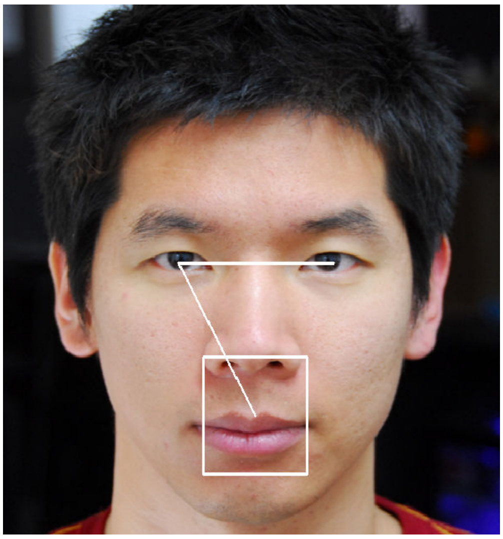


Fig. 8. Candidate region of lips

III.3. EXTRACTING FEATURES OF NOSE

After eyes and nose are found, when the characteristic of the nose in the candidate region receives light, the light intensity of the nostril is relatively low. Therefore the characteristic sampling of nose is achieved by the local thresholding of grey levels. In this method the nostrils are either not shown or shown as connected depending on the lighting. However since the lip center coordinative value is obtained, by finding one point relatively away from the lip center coordinative value, the remain point can be taken symmetrically from the center of nose. Figure 10 is the process of extracting nose.

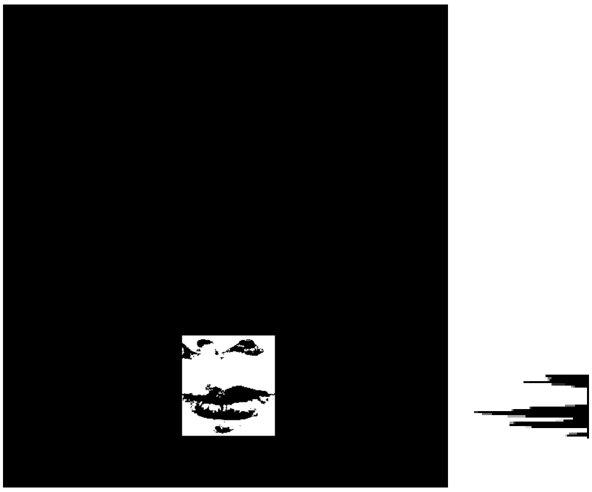
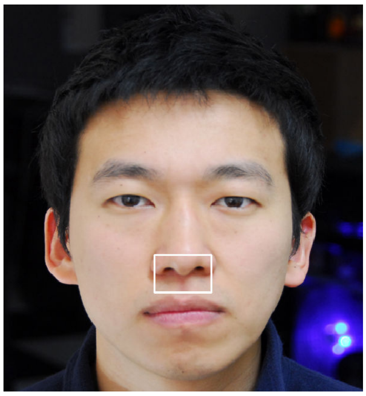
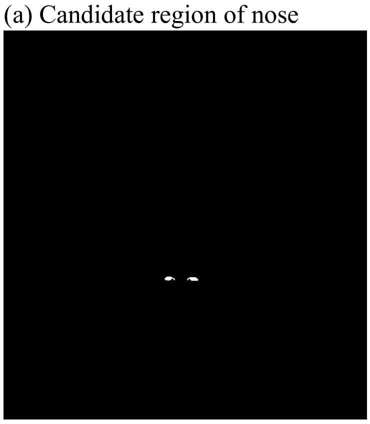


Fig. 9. Cumulative distribution of X axis in lip region using a*.



(a) Candidate region of nose



(b)Local thresholding image of nose

Fig. 10. Process of extracting feature of nose

III.2. **EXTRACTING FEATURES OF FACE OUTLINE**

We extracted p,q,r,s,t which is the points that straight line from a,l,o,n,g to end of skin region. The skin region is extracted by found average and deviation of YCbCr after drawing a square that based on ‘a’ and ‘g’, the points inside the square is the skin region. Figure 11 is extracted skin region to find feature points of outline of face. The point of "r" in the face is same skin region as neck, so it is hard to extract as a value of skin region, so judged by the meeting point with the point of lip center after extracting candidate region of "r' and extract edge. Heavily bearded face judged by the meeting point with skin region and this is the subject to be studied more.

Fig. 11. Extracting skin region

IV. EXPERIMENT

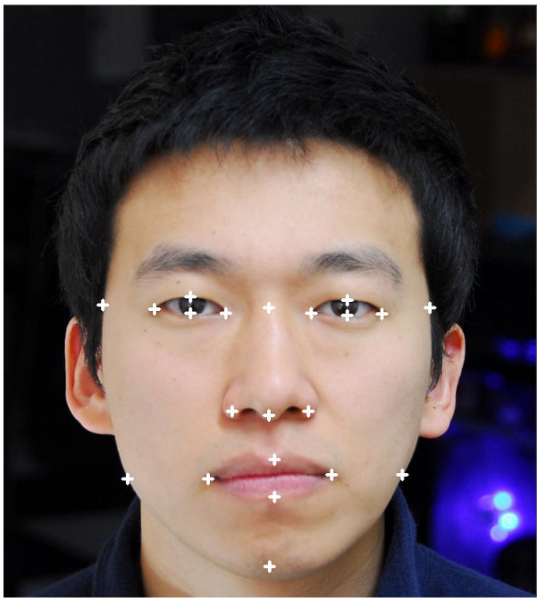
To perform this study, the face image database is largely divided into two. This is face images of 200 candidates from i-Phone of Apple company. They were taken at a comfortable office environment or outside before it gets too dark. If possible, photographers were asked to take many or take a picture of themselves. The limitation of the photograph was set to only show one face, if possible front face was taken. These photographs size were varied from 309 x 380 to 1892 x 2220. The collected images were mostly the yellow race, and in order to obtain the white and the black race 100 face images of various race were collected in the internet. The collected standard was one front face image similar to images taken from i-Phone and the resized image of 582 x 582 were used.

The study images were consisted of 100 of the yellow race, 50 of the white and 50 of the black. To test the robustness of algorithm, images of the same person in various lighting condition were also included, the test images were consisted of total of 180 different candidates. While collecting the experiment images, we found white race has relatively thin lips than yellow race and the space between eyebrows and eyes were closer than yellow race. Movie star Tom Cruise can be the example. Black race has very thick lips and relatively big nose than other races, movie star Denzel Washington may be the good example of this. In this experiment used CPU 2 GHz PC and OpenCV Library to increasing the speed except important algorithm. Processing speed was 0.789 seconds ~ 1.485 seconds per image. Table 1 is experiment result. Figure 9 is example of success image.

For the result of experiment, after we marked features which we wanted to shoot on the 200 pictures of experimental image, if we extract features by presented algorithm in 7x7 region from the value of marked features, we consider it as a success.

This is the error within 1% of total images and in actual shooting of people, there was difference somehow. Figure 1 is the result of experiment. If we analyze the failures, the eyes was stick to eyebrow so closely that mis-extracted the end point and the image of nose was lightened from below side so there was no difference of the light and shade between nostrils and other parts, and in the lip case it is too thick to measure thickness exactly. The cases of fail extracted, the lighting conditions were un-uniformity or the clearness was too low. In pre-treatment stage of this study, the algorithm of lighting equalization and Histogram equalization were applied to revise equalization and clearness but the algorithm of this part has to be studied more and in case of eyes and lips, the extraction shall be done more correctly thru physical correction.

Fig 12. Example of success image



V. CONCLUSION

Recently as face recognition process, not only the extraction of face features but also the change of lip shape and movement and for seeking the correct shape of eyes, nose and lips in order to convert 2D to 3D, studies are under proceeding lively.

In this paper we suggest you the way to extracting 20 characteristics from 2D face images and define 20

characteristics of face region to convert 2D face image to 3D image. In case of eyes, we found end points of eyes by extracting eye balls and calculate the thickness of eyes.
And then in case of lips we found end points after found thickness of lips by suggesting La*b* color space and set the exact candidate region. In case of nose we could get the end points and midpoint by setting the candidate region after compared to coordinate of midpoint of lips.
In this paper we tried to find eyes, nose and mouth by setting the exact candidate region and did many experiments by comparing several color spaces. In the future subject will find more detailed and exacted characteristics by using color space and character of shape.
When we look about the fail extraction the lighting conditions were not uniformity or the clearness was low or deep-set eyes or black man's lip, or bushy-bearded was not uniform. For the improvement of such cases, the study of lighting equalization algorithm must be done and the algorithm to extract jaw line exactly using morphology must be proposed. L*a*b color space will be studied more and the correctly extracted features will be applied to various practical fields.

Table 1. Experiment result

Recognition Region	Success number	Fail number	Success rate(%)
Eye	190	10	95
Nose	194	6	97
Lip	184	16	92
Face	190	10	95

REFERENCES

[1] E. Saber, A. M. Tekalp, "Frontal-view Face Detection and Facial Feature Extraction Using Color, Shape and Symmetry Based Cost Function," Pattern Recognition Letters 19, pp.669-680, 1998.
[2] S. Teskeridou, I. Pitas, "Facial Feature Extraction in Fontal Views Using Biometric Analogies," European Signal Processing Conference, Vol. I, pp.315-318, 1998
[3] J. Yang, R. Stiefelhagen, U. Meier, A. Waibel, "Real-time Face and Facial Feature Tracking and Application," Proc. Auditory-Visual Speech Processing, pp.79-84, 1998.
[4]A. R. Mirhossenei, H. .Yan, K. M. Lam, "Adaptive Deformable Model for Mouse Boundary Detection," Optical Engineering, Vol. 37 No. 3, pp.869-875, 1998.
[5] T. Wark, Sridharan, V. Chandran, "An Approach to Statistical Lip Modeling for Speaker Identification via Chromatic Feature Extraction," IEEE ICPR'98, Vol. 1, pp.123-125, 1998
[6] R. L. Hsu, Abdel-Mottaleb, M. Jain, A.K., "Face Detection in Color Images," IEEE trans. Pattern Analysis and Machine Intelligence, Vol. 24, pp.696-706, May 2002.
[7] N. Otus, "A Threshold Selection Method from Gray-level Histograms," IEEE Trans. System Man Cybernet. SMC-8, pp.62-66, 1978.
[8] J. S. Lee, Y. M. Kuo, P. C Chung, E. L. Chen, "Naked image detection based on adaptive and extensible skin color model," Pattern Recognition Society, vol.40, pp.2261-2270, 2007.
[9] P. Viola and M. J. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," IEEE CVPR. Vol.1. pp.511-518, 2001.
[10] R. Lienhart and J. Maydt, " An Extended Set of Harr-like Features for Rapid Object Detection," IEEE ICIP, pp.900-903, 2002
[11] Ming Husan Yang, David J. Kriegman, Narendra Ahuja, "Detecting Faces in Image: A Survey'" IEEE Transactions on PAMI, vol.24.No.!,Jan. 2002
[12] Ron Meirand and Gunnar Ratsch, "An Introduction to Boosting and Leveraging," Advanced Lectures on Machine Learning: 118-183,2003
[13]Robert E. Shcapire, Yoav Freund, Peter Bartlett, Wee Sun Lee, " Boosting the Margin : A New Explanation for the Effectiveness of voting Methods," The Annals of Statistics, 26(5): 1651-1686,1998
[14]P{aul Viola, Michael Jones, "Fast and Robust Classification using Asymmetric Adaboost and a Detector Cascade," Neural Information System 14, Dec. 2001.
[15]Michael J. Jones, Paul Viola, " Face Recognition Using Boosted Local Features," International Conferences on Computer Vision, 2003
[16]G. Yang and T.s. Huang, "Human Face Detection in Complex Bacground," Pattern Recognition, Vol.27, No.1.pp53-61,1994
[17]D. Chai and King N. Ngan, " Face Segmentation Using Skin-Color Map in Video Phone Applications," IEEE Trans. On Circuits and System for Video Technology. Vol. 9. No.4, pp.551-564, Jun. 1999.
[18]J. Kim, J.Zhu and H. Lee, "Block-Level Processing of a Video Object Segmentation Algorithm For Real-time System," in Proc. Of IEEE International Conference on Multimedia & Expo. Jul, 2007, pp.2066-2069
[19]Y. Deng, G. Su, K. Zhou, B. Fu, "Fast and Robust Face Detection in Video," in Proc. Of the fourth International Conference on Machine Learning and Cybemetics, Aug. 2005,, pp. 4577-4582.
[20] L. Craw, D. Tock, and A. Bennett, " Finding Face Features," Compuer Vision-ECCV'92, Lecture Notes in Computer Dcience, Vol.588, pp.92-96,1992.
[21]H. Joo, B. Jang, R. Bashar, and P.Rhee,"Real Time Face Detection Syatem Based Edge Restoration and Nested K-Means at Frontal View," Fuzzy Systems And Knowledge Discovery, Lecture Notes in Computer Science, Vol. 4223, pp.1192-1200,2006.
[22] K. C. Yow and R. Clipolla, " Detection of Human Faces under Scale, Orientation and Viewpoint Variations," Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, pp.295-300,1996.
[23]H.. Rowley, S. Baluja and T. Kaneda, "Neual Network Based Face Detection," IEEE. Tran. On PAMI, vol20, no.1.Jan. 1998.
[24]P. Jeonghee, S. Jungwon, A. Dongun, and C. Seongjong, "Detection of Human Faces Using Color and Eyes," in Proceeding of IEEE International Conference on Multimedia and Expo, vol.1. pp.133-136,2000.
[25]Shinjiro Kawato and Hun Ohya, "Automatic Skin Color Distribution Extraction for Face Detection and Tracking," in Proceeding of ICSP, pp.1415-1418,2000.
[26] C. Garcia and G. T. Ziritas,"Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," IEEE Transaction on PAMI, vol23. no.1, Jan. 2001.

Ha-sung Koo
Prof. Koo graduated from Kwangwoon University in Korea in 1989 with an Electrical Engineering degree. He received MS degree and Ph.D. form Kwangwoon University in Electrical Communication department in 1995. He is Associate Professor of Computer and Information Engineering at Hanseo University, and his research interests include machine vision, pattern recognition and 3D image processing.

Ho-Geun Song
Prof. Song received his B.S degree from Department of Electronic Engineering, ChungAng University, Korea, in 1991 and his M.S and Ph.D degrees from Graduate School of ChungAng University, Korea,in 1993 and 1997, respectively. Since 1996, he has been a faculty member of the Hanseo University in Korea, where he is now a professor at the Department of Computer and Information Engineering, and holding the position of Infomation Technology Officer at Information Technology Service Center in Hanseo University since 2007. His current research interests are Digital image processing, Contents-based image/video retreival, 3D image processiong, 3D object recognition, Digital image forengics, etc.