DICTA2002: Digital Image Computing Techniques and Applications, 21--22 January 2002, Melbourne, Australia.

Chin Extraction in Colour Frontal-view Face Images

Sakina Boursas, Hong Yan
School of Electrical and Information Engineering, University of Sydney, N.S.W. 2006 Australia
melissa@ee.usyd.edu.au, yan@ee.usyd.edu.au

Abstract

Among the various features of the face, the chin contour is usually estimated or approximated to some geometrical shapes. It is addressed in the present work for extraction. In our approach, a skin colour segmentation algorithm is initially implemented to roughly localise the face area. Because colour-based segmentation methods are too sensitive to lighting, region homogeneity is introduced to clearly and robustly separate the features of the face. Using the eyes' locations, an estimation of the chin region is made, and black top-hat morphological operation is then applied to extract the chin contour. The resulting set of discontinued segments and fragmented curves are processed, thinned then classified before being grouped. Some local interpolations can then be made, if needed, to reconstruct the entire chin contour. Experiments are carried out mainly on images from the Surrey Database to evaluate the robustness of the method. The results are very encouraging.

1. Introduction

Facial features detection plays an important role in a vast variety of human face-based applications: recognition of faces, facial expression analysis, face detection, identity verification, head pose estimation, orthodontic applications [1,2] and face and features tracking algorithms. One notices that, although a majority of facial features are included in many face-based applications, only a few features such as the eyes and mouth are frequently addressed for fine extraction.

In a psychological assessment on the cue saliency of facial features [3], the chin is ranked less important than the eyes but as important as the mouth and the lips; furthermore in anthropological studies, the shape of the chin and the head are used to determine race [4,5].

The chin was used for distance measurements in early human face recognition schemes [6,7].

For face recognition purposes, an attempt was made by Jia and Nixon [8] to extract the chin boundary using frontal and profile views of faces but the results were poor. Recently chin and cheeks contours attracted some interest in the video coding community but only estimation and approximations of the contours with curves and parabolas have been performed so far [9-11].

We introduce in the present work a new method for chin contour extraction and reconstruction. A more robust to lighting face segmentation algorithm is applied, allowing a clear separation of the facial features inside the face.

In our approach, the face area is first localised with a method combining two colour-based, pixel-level segmentation algorithms. The region homogeneity of the face is computed to segment the interior of the face area, making the face features distinct from the boundary. Some knowledge about the face structure is used to localise the eyes; then the chin region is estimated for local processing. Applying morphological black top hat operation proves to be effective for chin boundary extraction. The resulting segments and fragmented curves are thinned and classified. A number of consistent rules based on object geometry and curvature have been set to assemble the relevant parts of the chin. Missing parts in the chin contour are interpolated using the neighbouring segments.

The paper is organised as follows. Two approaches for face skin colour segmentation are presented in Sections 2.1 and 2.2 respectively, then a brief description of homogeneity and its implementation for segmentation of the face region are described in Section 3. The methodology used to estimate the chin area is explained in Section 4.1. Chin contour extraction using top hat morphological operation is shown in Section 4.2. Chin contour reconstruction is detailed in Section 5. The experiments and discussions are given in Section 6 and finally the conclusions are drawn in Sections 7.

2. Face skin-colour based segmentation

Skin colour-based segmentation methods are numerous and habitually used as an initial step in face based applications. Despite their popularity, their performance tends to diminish for varying lighting conditions causing false positive or negative detection, which can be misleading in a facial features localisation task. To alleviate some of these effects, we have chosen to combine two pixel-level skin colour detection methods, to locate the face area. The eyes are localised using homogeneity, as presented in section 3. The two skin colour-based approaches are chosen for their simplicity and low computational cost.

The first approach uses normalised ratios in RGB space and the second a linear transformation of YIQ space. We only review these approaches very briefly in sections 2.1 and 2.2, further details can be found in [12].

2.1. Normalised RGB ratios approach

Two large sets of skin and non-skin pixels were collected from Compaq Database [13] and plotted in the RG, RB and GB planes. The lines that separate the skin and non-skin pixels in the three planes were extracted visually from the plots. These limits represent the thresholds used to detect skin pixels in images.

The following set of inequalities is employed for skin segmentation:

$$L_{rg} < R/G < U_{rg}$$
$$L_{gb} < G/B < U_{gb}$$
$$L_{rb} < R/B < U_{rb}$$

R, G and B are the three colour components of the image to be segmented. L_{rg}, U_{rg}, L_{gb}, U_{gb}, L_{rb} and U_{rb} are the lower and upper limits deduced respectively from red versus green, green versus blue and red versus blue plots.

2.2. YIQ Colour space transformation

Wang [14] proposed a linear transformation of the YIQ colour space for the detection of skin pixels. The new colour space, YI'Q', is obtained by applying a rotation to the IQ plane with respect to the Y-axis. The matrix of transformation is calculated as follows:

When the angle of rotation q is carefully chosen, the projected skin colour pixels in I'Q' plane are constrained and separated from the non-skin pixels.

In the same manner as that in Section 2.1, the Compaq Database is used for threshold determination from the I' versus the Q' plot. In our algorithm an extension has been made to the Q' component. The following set of inequalities is used to select the skin pixel in the YI'Q' colour space:

$$R_{I}' < I' < L_{I}'$$
$$R_{Q}' < Q' < L_{Q}'$$

where $R_{I', Q'}$ and $L_{I', Q'}$ are respectively the right and left limits in the I'Q' and Q'I' planes.

The two Images ensuing from the skin-colour segmentation algorithms are combined with a logical 'And' for better results. Holes in the segmented head region are filled. Figure 1 illustrates the steps explained above.

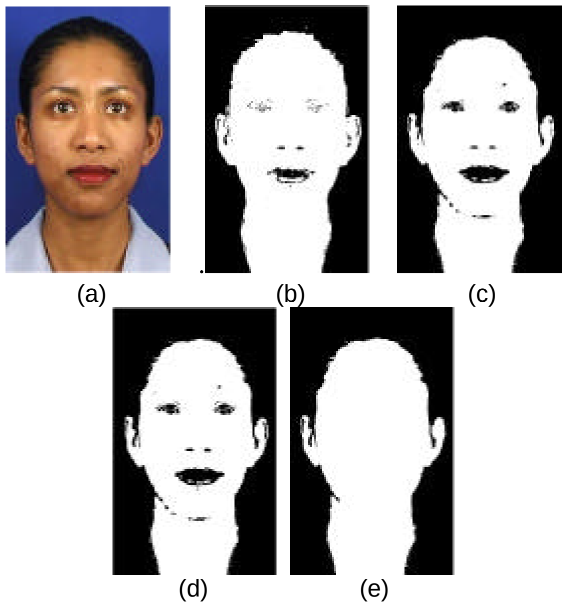


Figure 1. Skin colour-based segmentation. (a) Original image, (b) RGB-ratio based segmentation, (c) I'Q' based segmentation. (d) Combination of (b) and (c). (f) Segmented face region.

3. Region Homogeneity

Homogeneity is related to region uniformity. I it was introduced by Cheng and Sun [15] as a composition of the standard deviation representing the contrast within a region and the intensity discontinuity caused by an abrupt change in grey levels. Homogeneity is applied in our approach to simplify the image data while preserving the most relevant shapes and features. The homogeneity of the segmented face area has been computed from the grey-level image with the following formulae:

$$H(g_{ij}, v_{ij}^{(1)}, v_{ij}^{(2)}) = 1 - E(g_{ij}, v_{ij}^{(2)}) \times V(g_{ij}, v_{ij}^{(1)}) \quad (1)$$

where g_{ij} is the intensity value at location (i,j).

$v_{ij}^{(1)}$ and $v_{ij}^{(2)}$ are two windows of sizes $d \times d$ and $t \times t$ both centred at position (i,j).

$E(g_{ij}, v_{ij}^{(2)})$ and $V(g_{ij}, v_{ij}^{(1)})$ represent the normalised values of the standard deviation nij and the discontinuity, eij and are calculated as follows:

$$E(g_{ij}, v_{ij}^{(2)}) = \frac{e_{ij}}{e\max} \quad (2)$$

$$V(g_{ij}, w_{ij}^{(1)}) = \frac{v_{ij}}{v\max} \quad (3)$$

where the standard deviation is given as:

$$V_{ij} = \sqrt{\frac{1}{d^2} \sum_{p=i-(d-1)/2}^{i+(d-1)/2} \sum_{q=j-(d-1)/2}^{j+(d-1)/2} (g_{pq} - m_{ij})^2} \quad (4)$$

And the mean value over the grey-level values within the window $v_{ij}^{(1)}$ is given by:

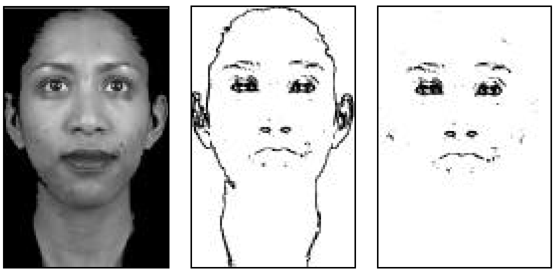
$$m_{ij} = \frac{1}{d^2} \sum_{p=i-(d-1)/2}^{i+(d-1)/2} \sum_{q=j-(d-1)/2}^{j+(d-1)/2} g_{pq} \quad (5)$$

The discontinuity value is obtained with Sobel operator:

$$e_{ij} = \sqrt{G_x^2 + G_y^2} \quad (6)$$

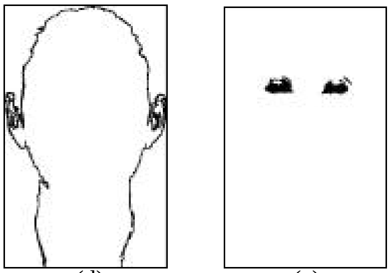
G_x and G_y are respectively the gradient magnitudes in the x and y directions.

Calculated as described above, the homogeneity at each pixel in the image ranges from zero to one. Since we do not perform colour segmentation as presented in [15], we proceed to a thresholding of the image by setting a high homogeneity threshold for a clear separation between the features and the face boundary. See Fig 2.



(a) (b) (c)

Figure 2. Homogeneity-based face segmentation. (a) Grey-level image of face area. (b) Homogeneity of the bounded face area after thresholding. (c) Extracted boundary of the face region.



(d) (e)

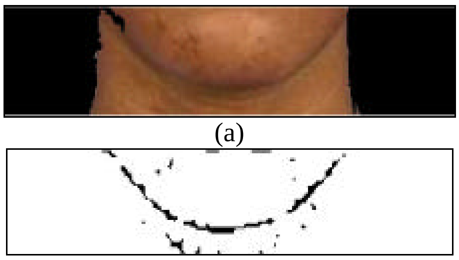
Figure 2. (Continued) Homogeneity-based face segmentation. (d) Features remaining after boundary extraction (e) Localised eyes.

4. Chin Contour Extraction

4.1. Chin Area Estimation

To perform a local processing on the chin, an estimation of its region, delimited by the upper limit Chin_lim1 and the lower limit Chin_lim2, is first made. The dimensions of the face region are not considered due to false detection introduced in the colour segmentation, instead the horizontal position of the eyes [16] is used to compute the values Chin_lim1 and Chin_lim2.

To localise the eyes, the head contour is eliminated as follows, the head area obtained after homogeneity segmentation is enclosed in a bounding rectangle. All the connected components of the binary image that touch the box frame according to the connectivity of a 3×3cross structuring element are extracted. In the remaining picture, small objects are removed as well as long and thin vertical segments often representing hair falsely detected as skin pixels. Once the holes are filled, a search for a pair of compact objects horizontally aligned is initiated to locate the eyes; subsequently horizontal lines at Chin_lim1 and Chin_lim2 are determined.



(a)

(b)

Figure 3. (a) Chin area extracted. (b) Black top-hat applied on the chin region.

4.2. Top hat for Chin Boundary Extraction

The chin is a feature that is not characterised by a particular colour or shape, as are the eyes and the mouth. Using Black top hat operation to address chin contour extraction has been suggested by the presence, in the chin region, of shading due to the inherent structure of the face. To enhance the details of the chin contour, a 3×3 circular structuring element is selected. Broken segments and fragmented curves resulting from the grey level top hat operation are then processed for reconstruction.

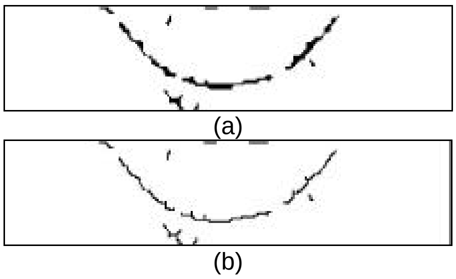


Figure 4. Contour. (a) Outcome of black top hat operation cleaned. (b) Thinned segments.

5. Chin Contour Reconstruction

The reconstruction of the chin contour is exclusively performed with the outcome from black top hat operation. We do not make any assumption on the shape of the chin nor do we use approximations to its contour. For the selection of relevant chin segments, objects with a low compactness value are first removed. Objects, which are candidates for chin reconstruction, should have a high compactness value i.e. more pixels to form the perimeter than to fill the area. Remaining segments are thinned then classified according to their length. Long segments are partitioned in smaller sub-segments to compute their curvature while short ones are kept unchanged. Segments are then sorted according to their point of occurrence. The slopes and slope of connections made between separated segments are analysed by looking globally at the pixel values, slopes and connections.

The curvature of the grouped segments selected for reconstruction should first be negative then monotone positive. Neighbouring segments to gaps are used for local interpolation for the reconstruction of the entire chin contour (see Fig5 (b)).

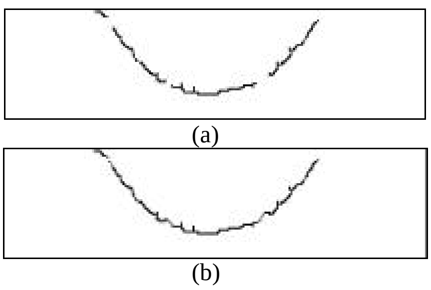


Figure 5. (a) Grouped chin segments. (b) Reconstructed chin contour.

6. Experiments and Discussion

Experiments were performed on frontal-view face images from the Surrey database [17]. Subjects from various backgrounds and different chin shapes were selected to access the performance of the proposed algorithm. Images were cropped to allow faster processing.

The algorithm implemented for chin extraction and reconstruction performs notably well, with the extraction section performing better than the reconstruction one. The main difficulty in the reconstruction part being the selection and grouping of the relevant scattered pieces of the chin. From the results, it is clear that the use of homogeneity has minimised the effects of lighting and considerably helped to locate the features even in the presence of hair. (see (Figure 6 (Subject 9))).

The chin was correctly extracted and reconstructed in most pictures, reaching the boundary of the face. While the face boundary is a useful feature for the reconstruction of the chin, it was not included because of the imperfections introduced by colour segmentation methods in still images. A “Chubby chin” was successfully extracted in Figure 6 (Subject 4). It should be noted however that the case of superimposed chins was addressed during implementation. In similar cases the upper contour is considered to be the chin contour. In Figure 6 (Subject 8), the shading caused by the presence of a scar on the lower part of the chin has altered the extracted contour of the chin slightly.

7. Conclusions

A new method was implemented which successfully extracts the contour of the chin. Faces are segmented using colour-based methods and homogeneity. The chin area is extracted using the eyes’ position. Extraction is performed effectively with the black top hat operation, providing great precision on the chin contour. The results are very encouraging and could be useful for many face-based applications. The proposed top hat algorithm for chin extraction provides a better description of the face shape than any geometrical or mathematical model proposed so far. For recognition purposes this method is more beneficial.

Introducing homogeneity has alleviated some of the negative effects of lighting on the selected set of pictures, but for multiple faces and complex background images more robust methods should be investigated. For a better estimation of the chin area other features than the eyes positions could be used. We are currently working on an adaptive colour-based segmentation method and on improving the chin reconstruction algorithm to implement the method for more complex and poorer quality colour frontal-view face images.

Acknowledgments

This work is supported by the Australian Research Council.

We have made use of the Extended Multimodal Face Database and associated documentation. Further details of this software can be found in; K. Messer, J. Mats, J. Kittler, J. Luettin and G. Maitre; “XM2TVS Database, Proceedings 2nd Conference on Audio and Video-base Biometric Personal Verification (AVBPA99) Springer Verlag”, New York, 1999. CVSSP http://www.ee.surrey.ac.uk/Reaserch/VSSP/xm2vtsdb”

We would like also to thank Compaq for the use of their skin and non-skin database [14].

References

[1] A.M. Coombes, R. Richards, and et al., "Shape-based description of the facial surface." In IEE Colloquium on Machine Storage and Recognition of Faces, London, UK, January 1992, pp 9/1--9/4.

[2] M.A. O'Neill, M.I. Denos, and et al. "An automated technique for the generation of facial surface models as an aid in orthodontic and facial orthognathic research0." In IEE Colloquium on Machine Storage and Recognition of Faces, London, UK, January 1992, pp 11/1--11/6.

[3] G. Davies, H. Ellis, and J.Sheferd, Perceiving and remembering faces, Press Series in Cognition and Perception.,: Academic Press, London, 1981.

[4] Stephen Molnar, Human Variation: Races, Types, and Ethnic Groups, Fourth Edition. Englewood Cliffs, New Jersey: Prentice Hall. 1998

[5] Gill et al. in "Craniofacial Variation in Forensic Race Identification," (IN: Forensic osteology: advances in the identification of human. Kathleen J. Reichs (ed.); [Springfield, Ill., U.S.A.: Thomas, 1986.

[6] T. Sakai, M. Nagao, and T. Kanade, "Computer analysis and classification of photographs of human faces.",Proc. First USA-Japan Computer conf., 1986, pp. 245-256.

[7] R. Brunelli and T.Poggio, "Face Recognition: Features versus templates", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 10, Oct. 1993, pp 1042-1052.

[8] Xiaoguang Jia and Mark S. Nixon, " Extending the Feature Vector for Automatic Face Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 12, December 1995, pp 1167-1176.

[9] M. Kampmann, "Estimation of the chin and cheek contours for precise face model adaptation ", International Conference on Image Processing (ICIP--97), Santa Barbara, USA, October 1997, pp. 300--303.

[10] M. Kampmann, "Precise face model adaptation for semantic coding of videophone sequences", Picture Coding Symposium, Berlin, Germany, Sept. 1997.

[11] M.Kampmann: "Segmentation of a Head into Face, Ears, Neck and Hair for Knowledge-Based Analysis-Synthesis Coding of Videophone Sequences" International Conference on Image Processing (ICIP98), Chicago, USA, October 1998.

[12] Jason Brand and John S. Mason, "A Comparative Assessment of Three Approaches to Pixel-level Human Skin-Detection.", International Conference on Pattern Recognition, vol. 1, 2000, pp.1056-1059.

[13] M.Jones and J. Rehg., "Compaq Skin Database". http://www.crl.research.digital.com/publications/tech-reports/abstracts/98-11.html

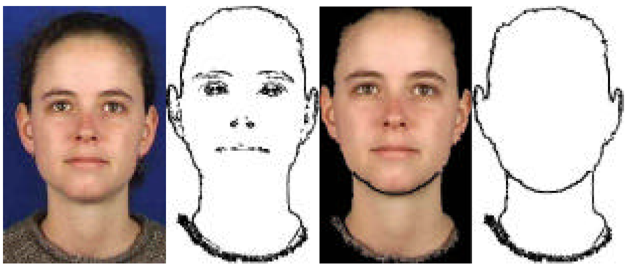
[14] Wang and M.Brandstein, "Multi-source face tracking with audio and visual data". IEEE MMSP, 1999, pp168.

[15] Heng Cheng, Ying Sun, "A Hierarchical approach to colour image segmentation using homogeneity.". IEEE Transactions on Image Processing, Vol. 9, No12, December 2000.

[16] Alattar and S. A. Rajala, "Facial features localisation in front view head and shoulders images", IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 6, 1999, pp. 3557--3560

[17] The Surrey Database http://www.ee.surrey.ac.uk/Reaserch/VSSP/xm2vtsdb

Subject 1



Subject 2

(a) (b) (c) (d)

Figure 6. Chin Extraction and Reconstruction Results. (a) Original image (b) Homogeneity based segmentation (c) Extracted reconstructed chin (c) Entire Boundary of the face.

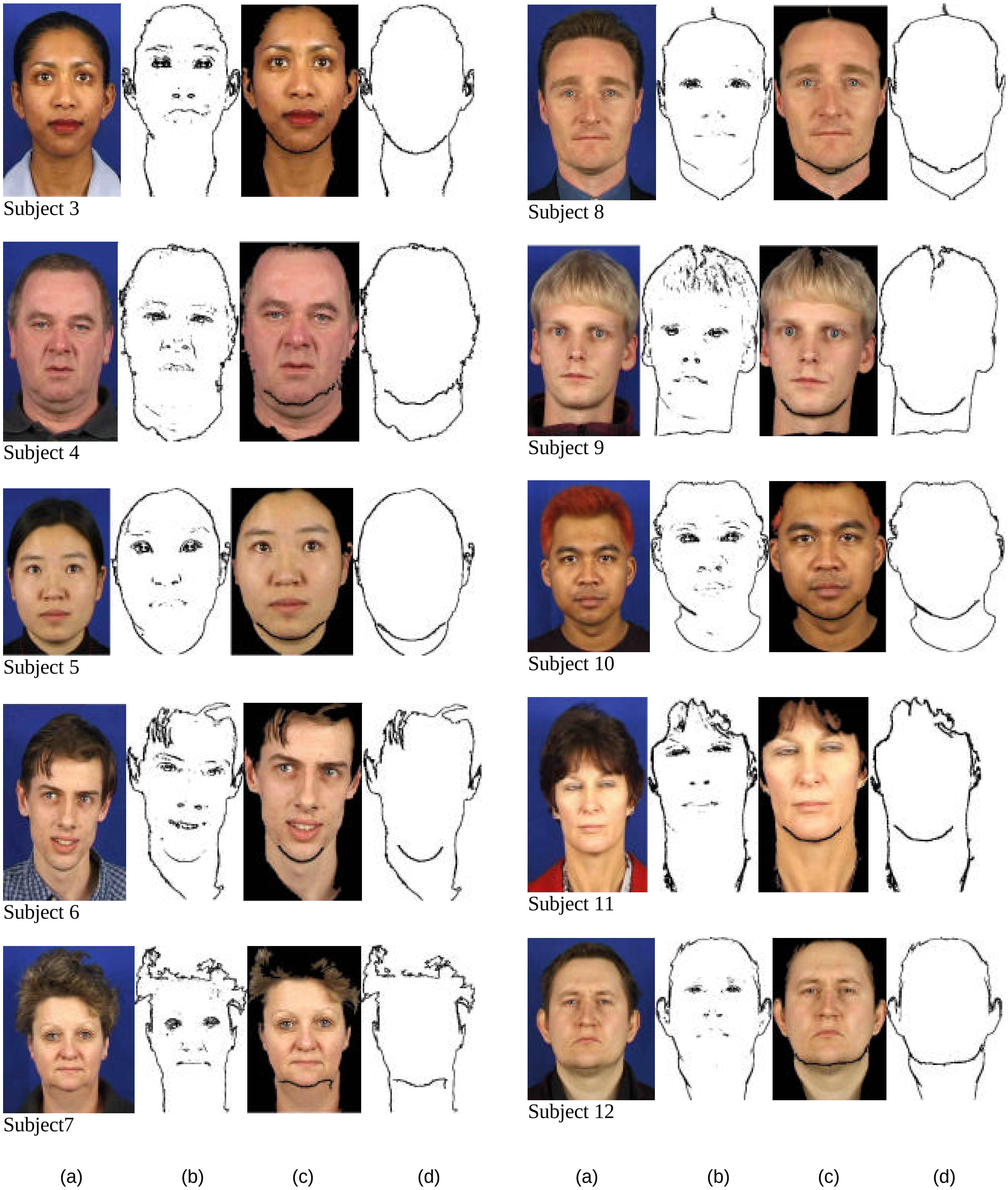


Figure 6. (Continued) Chin Extraction and Reconstruction Results. (a) Original image (b) Homogeneity based segmentation (c) Extracted reconstructed chin (c) Entire Boundary of the face.