Feature Extraction and Localisation using Scale-Invariant Feature Transform on 2.5D Image

[1]SukTing Pui, [2]Jacey-Lynn Minoi, [3]Terrin Lim
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak
94300 Kota Samarahan, Sarawak, Malaysia
[1]polly_sukting@hotmail.com, [2]jacey@fit.unimas.my, [3]terrin@fit.unimas.my

João Fradinho Oliveira
C3i/Instituto Politécnico de Portalegre, Portalegre
Portugal
jfoliveira@estgp.pt

Duncan Fyfe Gillies
Department of Computing
Imperial College London
United Kingdom
d.gillies@imperial.ac.uk

ABSTRACT
The standard starting point for the extraction of information from human face image data is the detection of key anatomical landmarks, which is a vital initial stage for several applications, such as face recognition, facial analysis and synthesis. Locating facial landmarks in images is an important task in image processing and detecting it automatically still remains challenging. The appearance of facial landmarks may vary tremendously due to facial variations. Detecting and extracting landmarks from raw face data is usually done manually by trained and experienced scientists or clinicians, and the landmarking is a laborious process. Hence, we aim to develop methods to automate as much as possible the process of landmarking facial features. In this paper, we present and discuss our new automatic landmarking method on face data using 2.5-dimensional (2.5D) range images. We applied the Scale-invariant Feature Transform (SIFT) method to extract feature vectors and the Otsu's method to obtain a general threshold value for landmark localisation. We have also developed an interactive tool to ease the visualisation of the overall landmarking process. The interactive visualization tool has a function which allows users to adjust and explore the threshold values for further analysis, thus enabling one to determine the threshold values for the detection and extraction of important keypoints or/and regions of facial features that are suitable to be used later automatically with new datasets with the same controlled lighting and pose restrictions. We measured the accuracy of the automatic landmarking versus manual landmarking and found the differences to be marginal. This paper describes our own implementation of the SIFT and Otsu's algorithms, analyzes the results of the landmark detection, and highlights future work.

Keywords
Feature extraction, localization, landmarking, Otsu's algorithm

1. INTRODUCTION
The human vision system can perceive features such as the edges, tips or corners of an object, without any difficulties. For example, a human is able to detect and recognize the eyes, the tip of the nose and/or the mouth of a person at first glance. However, a computer is unable to do such task easily and effortlessly [CUS13]. The human vision system and brain mechanisms that are responsible for the detection of features are so complex that despite the work of neurobiologists, mathematicians and computer scientists, it is still not possible to replicate facial detection accurately.

In this paper, we propose methods to obtain distinct features on a face and label them automatically by placing landmarks. Our objective is to first empirically find specific landmarking threshold values that are valid for a given set of example datasets. Later, we use these found values in an automatic landmark setting with new example datasets and compare the accuracy versus manual landmarking.

Extracting facial information automatically is a challenging process, due to linear and nonlinear transformations, and therefore may not give a valid representation of the objects. The extents of spatial features differ significantly from different scales, sizes and resolutions. Therefore, the feature extraction methods ought to be invariant of scale and orientation. The SIFT approach takes a face image and transforms it into a collection of feature vectors. These features can either be global or local, defining the whole or a part of the image respectively. Usually the local interest descriptors which can be used to find distinct features on the face are selected manually. In our project, we make use of all the global and local descriptors to identify distinct features. These feature descriptors are usually in regions instead of at localised points. In order to localise these feature regions to an exact pixel point, a robust localisation algorithm is proposed based on all local maxima or minima, and the neighbouring pixels.

Identifying anatomical face landmarks, such as the corners of the eyes, the tip of the nose, the end corners of the lips, chin, etc., is a much easier task to those who are familiar with the anatomy of the face. On the contrary, extracting those landmarks automatically without the intervention of a human is a more complex process. Having to accurately identify a landmark point with minimal error is still a challenging problem and until today, there is no published method that is able to automatically extract the facial features and localise them to exact pixel points. This detailed information is crucial and is very important for applications such as surgical repair to improve facial appearance from cleft deformity, or other orthognatic surgeries, in which errors in landmarkings can cause serious problems. Specialist surgeons who treats hundreds of people with the cleft condition requires this detailed information.

A face image contains complex features with a large degree of background and face variations, such as the identity of the person, facial expression, head pose, facial hair variability, age, gender, cluttered background, etc. Face images also come with different format types, sizes, scales and rotation. All these variants lead to the difficulties in the automatic landmarking processes. The Scale-invariant Feature Transform (SIFT) was proposed to overcome these difficulties. SIFT was first introduced by David Lowe [Low04] and is an algorithm to detect and extract keypoints or local features in images. SIFT holds the advantage of being able to detect features even under changes in image scale, rotation and illumination. In contrast, in our work, rather than applying SIFT to regular 2D images, we adapt it to work on richer 2.5D images captured with a stereo camera system.

Over the last decades, 3D images have become popular due to the advancement of 3D sensor and camera technologies; alongside with 2.5D range images. Range images have a number of added advantages over 2D images. A 2.5D image is defined as a simplified three-dimensional (x, y, z) surface representation that contains at most one depth (z) value for every point in the (x, y) plane [Vau11]. One can think of a 2.5D image as a grey-scale image, where a black pixel corresponds to the background, while the white pixel represents the surface point that is nearest to the camera [ANRS07]. The 2.5D face images enable depth perception and allow one to manipulate the image like the 3D image. In addition to range data, colour perception on a face image is also possible. The various sources of information from the range image can be extracted to help derive different features regions. 2.5D range images are thus used in our system as a dataset to define the keypoint descriptors by extracting the facial surface information.

Our automatic landmarking consists of a generalised two-stage pipeline. In the first stage, feature extraction is performed on a 3D scale space volume using the state-of-the-art of robust and reproducible feature extraction techniques. In the second stage, we compute the weighted average of the extracted feature vectors to locate the centre or centroid of component feature regions. The coordinates of the centroid are given by the mean of the region at *x*- and *y*- coordinate.

In the first stage, Gaussian Smoothing is applied after Difference of the Gaussian (DoG) in order to extract weighted regions based on specific threshold values. Scale-space is constructed by taking the DoG of images at different scales. Then, within the scale-space, the weighted average of the curvature values of the elements is computed to estimate the centres or centroids. These regions of Mean (H) and Gaussian (K) curvatures [BJG85] map are coloured coded to ease visualization during the feature extraction stage. Next, the Gaussian Pyramid method is applied and the “Reduce” and “Expand” operations are employed at the successive pyramid levels to acquire filtered regions.

In the localisation stage, binary conversion is computed on the extracted features and the minimum threshold size of the local pixel area is set for landmarking. These stages will be applied in loops until the candidate landmarks and/or best regions are obtained. Otsu’s algorithm was applied to convert an image to a purely binary image by calculating a threshold to split the pixels into two classes. Otsu’s thresholding chooses a threshold to minimize the weighted within-class variance and maximize the between-class variance of the thresholded black and white.

We have developed an interactive Graphical User Interface (GUI) tool to ease the visualisation of the landmarking processes and to allow the manipulation of the extracted facial features. This tool allows one to select any of the nine surface primitives and thresholds can be adjusted to best get localised surface areas in order to identify suitable generalized threshold values.

The outline of this paper is as follows: In Section 2, we describe previous work related with our approach. In Section 3, we present our automatic landmarking method. In Section 4, the results and discussions are presented. Finally, Section V offers conclusions and directions for future research.

2. RELATED WORKS

In recent years, there have been a number of research approaches on face landmarking with automatic systems in different applications. Automatic landmarking has the potential to be very beneficial in areas ranging from face registration to facial expression recognition [SAC09]. Methods for landmarking on face models can be categorised in various ways. For example, the type or the modality of the data e.g a still image, range image, 3D data or even video sequence. The prior information underlying the sources determines the methodology to be used. In addition, an initial candidate facial landmarks selection and registration step is necessary in almost any application. Therefore, registration based on facial landmarks correspondence is the most crucial step to make a system fully automatic [PTK10].

Nowadays, the most common and popular method for face images analysis and processing relies on curvatures of their facial features [Sze12]. This approach has been widely used since the 1980s, whereby Besl and Jain [BJG85] introduced Mean (H) and Gaussian (K) curvatures to segment facial surface into eight different types. In early works, the following authors [TF95, Gor92, TIC98, KKK01, MAVD03] adopted this method and experimented with it on range images. Since then research has been carried out on how to represent the human face by applying features based on the shape and curvature of the face surfaces.

In addition, there are some other works on the Shape Index (SI) method, which is defined by Koenderick and van Doorn [KD92], which decouples the shape and the magnitude of the curvedness. In a series of publications reported in [CSJ05, JJC06, Col06], the authors presented methods to locate the positions of the corners of the eyes and mouth, the nose and also the tip of the chin. They have also developed a heuristic method to identify the tip of the nose more efficiently. The candidate landmarks were filtered by using a statistical model of landmark positions.

Nair and Cavallaro [NC09] presented a method to detect candidate facial landmarks on 2.5D scans. They applied the shape index and the curvedness index to extract feature points, the nose tip and eye corners. The feature points are fitted to the dataset according to three selected control points (nose tip and left and right inner eye corners) for the registration. However, the method is not applicable to pose self-occlusion cases, where missing data is not captured when using three control points fitted for the face models.

In [PPTTK09, PTPK09], the authors presented methods for better detecting facial landmarks on 2.5D scans. The candidate landmarks are divided into eye inner and outer corners, mouth corners, the nose tip and chin tips. In order to locate the candidate landmarks, local shape and curvature analysis is carried out applying shape index, extrusion maps and spin images. Landmarks are identified by matching them with a statistical facial landmark model.

Yu and Tiddeman [YT08] implemented the SIFT algorithm on 2D face images to find distinctive features in a sequence of face images. Their approach appeared to deliver distinct face features useful for face recognition. Gupta et al. [GKST10] employed the SIFT method for 2D face recognition on face images. They have extended the SIFT method by devising and adding probabilistic graphs to match SIFT keypoint features on independent face areas. The images used are controlled in illumination, scale and orientation changes. Each of the extracted features is a node and the relationships between invariant points are a geometric distance between nodes, thus a probabilistic graph. The sub images corresponding to the face limits are localised to eliminate noise.

The SIFT method can be extended to 2.5D with extra processing steps. Results have shown and demonstrated that the SIFT method is still effective in extracting features on 2.5D images. Guo et al. [GZJ10] applied SIFT to 2.5D range face images. Mean (H) and Gaussian (K) curvatures are combined with the Shape Index (SI) to extract facial features. Each keypoint can be divided into nine different primitive surfaces for a match. The results achieved are good and have robust rotation invariance. Cui et al. [CLDC12] constructed registration and integration algorithms for structured light 3D scanning using SIFT matching of multi-source images. This approach used 2.5D range images. They calculated the boundary of the overlapping regions, which is generated from the integration, and then identify landmarks. However, the process of obtaining grey information is not stable as there are certain depth changes and light reflections influencing the landmarks on the image. Han et al. [HYL08] employed SIFT on the human face to extract features for recognition. Images are tested under controlled conditions whereby the pose and facial expression are strictly limited. The experiment has shown high performances and robustness in facial expression recognition, and SIFT has claimed its invariance also towards illumination, noise, rotation and transformation.

 ISBN 978-80-86943-71-8

The following section presents our method; the extraction of features and how automatic landmarking is made possible.

3. METHOD

The 2.5D face range was captured with a commercial 3D face imaging system consisting of two SLR cameras positioned on each side. The camera system is used to produce a 3D surface mesh by automatically merging the acquired images. Only one dynamic range camera on the left hand side captures a 2.5D range image. Hence, the face is captured towards reference points positioned roughly at 45 degree angle to the left. Since the cameras are a commercial system, the details of the method are currently not available to the public. Two range images were acquired as a pilot experiment which were then used to conduct the automatic landmarking. No pre-processing method was performed while acquiring the range images.

The overall process of automatic landmarking can be generalised in a two-stage pipeline, as shown in Figure 1.

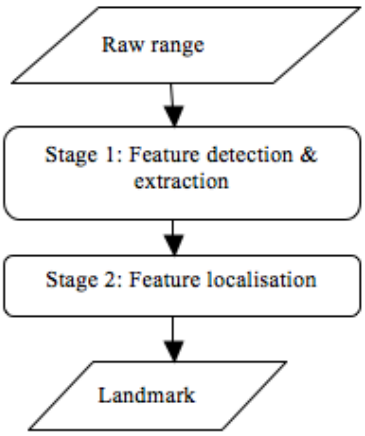

Figure 1. Two-stage pipeline landmarking.

Stage 1 feature detection and extraction is based on using SIFT. Its aim is to detect and extract facial features in the form of points, curves or regions. Once the features have been detected, patches around the features can be extracted. These features are known as feature descriptors or feature vectors. The methods used are Gaussian smoothing, binomial smoothing, Mean (H) & Gaussian (K) curvatures, and Gaussian of differences and Gaussian pyramid with 'Reduce' and 'Expand' operations within the layers of the pyramid. These functions performs the down-sampling and up-sampling steps of the Gaussian pyramiding construction.

Gaussian smoothing is applied to filter lower contrast and subsample images to provide a multi-scale representation of the image. Binomial smoothing is also applied and partial derivative estimate images were computed via the appropriate 2D image convolutions. This process is to search and determine the locations and scales, which are repetitively assigned under differing views of the same image. The identification of the keypoints can be accomplished by searching across all possible scales. Scale-space is constructed by taking the Difference of the Gaussian (DoG) images at different scale. Interpolation of nearby pixels from the DoG images is used to accurately determine the positions of each candidate keypoint. The keypoints with low contrast are removed and responses along edges are eliminated. Each pixel is examined by comparing with its eight neighbors at the same scale in DoG images and the nine corresponding neighbors at neighboring scales in 3x3 regions. A candidate keypoint is selected if the pixel is a local maximum or local minimum. The properties of the keypoint are measured with respect to the keypoint orientation, which provides rotation invariance.

In order to compute a feature descriptor which contains the signature of each of the main landmarks, it is important to identify the underlying primitive surfaces/patches and their surrounding pixels. In this stage, there are two algorithms applied - (i) Mean (H) & Gaussian (K) curvatures and (ii) Gaussian Pyramid - to determine the primitive surfaces.

The regions of Mean (H) & Gaussian (K) curvatures map are coloured coded (in RGB) to ease visualisation. The H&K curvatures are invariant of scale/resolution and orientation, whereby spaces are constructed in order to classify primitive surfaces into types [GZJ10]. There are nine types of primitive surfaces: peak, ridge, saddle ridge, flat, minimal, pit, valley, saddle valley and none. As shown in Table 1, each is labelled with a RGB colour.

The Gaussian pyramid method involved creating a series of images that is repeatedly convolved using Gaussian kernel to generate a sequence of images that are reduced or expand in scale. In our project, we construct images at decreasing scales then each of the image is convolved. From the original image, Level 0, the sequence of the convolved and reduced in scale images, Level 1, Level 2 etc are generated. These convolved images are copies of the original image reduced in scale by a factor of 2. Figure 2 shows an example of the levels of Gaussian pyramid method reduced in half the size of the original image. The left most image is Level 1 to Level 3 on the right most of the image.

It can be seen that at the higher levels of the pyramid, the smaller surface elements vanished and bigger elements reside. Afterwards, we expand the higher levels of the H&K pyramid to the original

 ISBN 978-80-86943-71-8

size. The details of Gaussian pyramiding can be found in [AABBO84].

At each level of the pyramid, 'Reduce' and 'Expand' operations are performed. In the 'Reduce' operation, the elements of the image are halved in both resolution and successive scales. The 'Expand' operation is computed to widen the primitive surface labels at each level of H&K pyramid to the original size.

	K>0	K=0	K<0
H<0 RGB	Peak T=1 150 150 150 (Grey)	Ridge T=2 255 0 0 (Red)	Saddle Ridge T=3 255 255 0 (Yellow)
H=0 RGB	None T=4 255 255 255	Flat T=5 0 0 255 (Blue)	Minimal T=6 0 255 0 (Green)
H>0 RGB	Pit T=7 255 255 255 (White)	Valley T=8 0 255 255 (Cyan)	Saddle Valley T=9 255 0 255 (Magenta)

Table 1. The details of H&K primitive surfaces and the colour labels used in the developed automatic landmarking tool [BJ88].

On each scale, H&K maps are computed. A new pyramid of H&K maps is obtained and used at every scale. Figure 3 shows the Level 3 of the pyramid to the size of the original image.

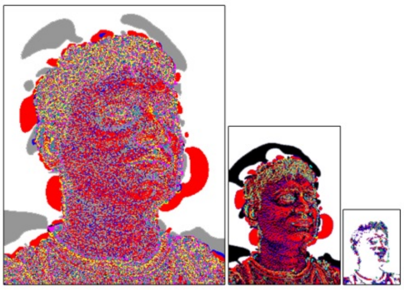

Figure 2. The Gaussian pyramid of the H&K maps.

Figure 3. Level 3 of the pyramid expanded to the size of the original image.

In this process, the primitive surfaces are examined to generalize threshold values by bootstrapping and resampling DoG images.

The next stage is the feature localisation process. It computes and places landmarks on the extracted primitive surfaces from the H&K maps after Otsu's method. Otsu's method [Ots75] calculates a global threshold value by calculating the spread within each of the classes. The aim is to minimize the weighted within-class variance and maximize the between-class variance of the thresholded black and white. The details of this algorithm can be found in [Ots75].

The foreground (the extracted primitive surfaces) will be converted to black pixels and the background is converted to white pixels. The extracted primitive surfaces are from the H&K map in Level 3 of the pyramid, for example as shown in the right most image in Figure 2 or the expanded size of the original image in Figure 3. The Otsu's thresholding is performed automatically to find a suitable general threshold value. The general threshold value corresponds to the valley of the historgram.

After Otsu's thresholding, the black pixel areas/regions in the thresholded image are evaluated based on the size of the local pixel areas to select the most suitable minimum and maximum threshold sizes. This can be done by firstly adjusting the threshold size values through our visualisation tool (as in Figure 8) and then by computing the weighted average of the local region to extract the centroid or mean of the local region/area. If the generated landmarks are undetectable or missing, the threshold size will need to be reduced.

Lastly, landmarking and labelling are computed on the thresholded image. Figure 4 illustrates the overview of the automatic landmarking processing.

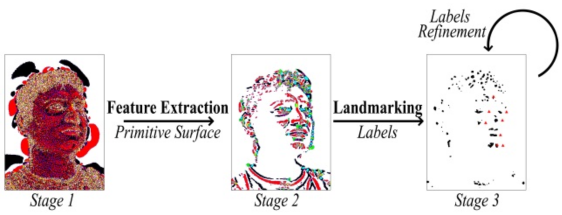

Figure 4. The overview of the automatic landmark process.

4. RESULTS AND DISCUSSIONS

Our Interactive Graphical User Interface (GUI) tool is presented Figure 8. It consists of a set of input elements that allows the adjustment of the threshold values during the landmarking process. Input boxes and selection buttons are equipped to perform the selection of the primitive surfaces for feature extraction and feature localisation.

There are nine keypoint descriptors being extracted, and each is examined by a set of minimum local pixel sizes for landmark labels. There are four sufficient extracted features regions namely peak, ridge, valley and flat (see Figure 5). Each of the extracted features image is processed with binary conversion in order to execute the landmarking process.

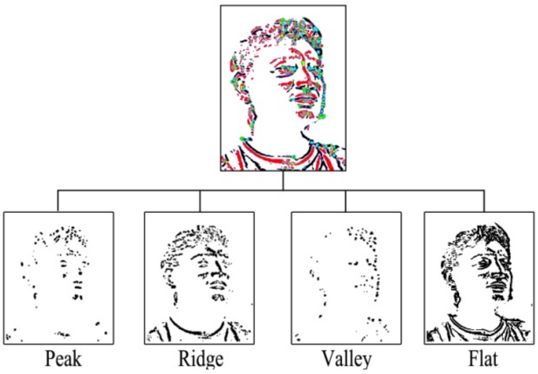

Figure 5. The extracted H&K features and converted to binary images.

Then, we execute localization process on these surfaces. We have identified from our results that the sufficient threshold values are ranging from 30, 40, 50 and 60 local pixel sizes. These values are tested accordingly and repeatedly until the highest accuracy of landmarks is obtained. From the analysis, we set the default pixel size to 40. The examples of identified landmarks on the extracted primitive surfaces are shown in Figure 9. Figure 9(c) shows that landmarks were successfully placed on top of the nose and the chin, while the others were unsuccessful. If the pixel size is set to 30 and below, the landmarks are mainly not in the face region and labeled on the hair and shirt. And if the value is set to 60 and above, the number of landmarks are limited on non-distinct features.

From our experiment, we have identified three primitive surfaces, namely peak, ridge and flat, that enable robust landmarking results on facial features. When tested on the other remaining datasets, we could also detect the nose tip and the chin automatically (see Figure 6 & 7). At the current stage of our work, we focus only on detecting distinct facial landmarks on the face region.

The automatic landmarking technique was compared to three sets of manual landmarking coordinates on the same dataset by three experts. The coordinates of the tip of the nose and chin were registered to perform an accuracy test on the automatic method.

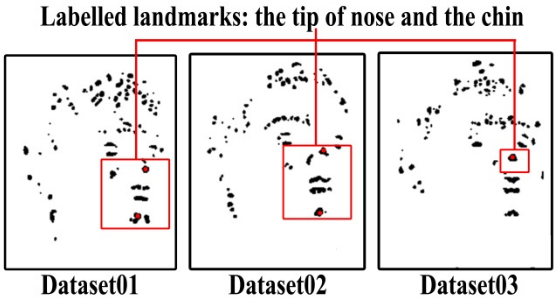

Figure 6. The landmark labeled on the tip of the nose and chin

The Root Mean Squared Error (Difference) was calculated between the data points in the automatic method with each manual registration points. Then, the accuracy formula given below was computed:

$$Accuracy, A = 2.4477x0.5\times\left(RMSE_x + RMSE_y\right)$$

Based on the results in Table 2, the automatic landmarking accuracy was divided into the two different regions. The automatic method performs relatively well for the detecting the tip of the nose and the chin, with low error and high accuracy. The accuracy is obtained by direct comparison of pixel with the automatic and manual landmark points. The average accuracy pixel differences for the tip of the nose is by 0.956 pixel while the chin is by 1.739 pixel.

	RMSE (x and y-coords)			
	Auto vs. Manual 1	Auto vs. Manual 2	Auto vs. Manual 3	Average Accuracy
Nose	x: 0.245 y: 0.719	x: 0.238 y: 0.408	x: 0.560 y: 0.173	
Accuracy, A	1.180	0.790	0.897	**0.956**
Chin	x: 1.298 y: 1.070	x: 0.825 y: 0.412	x: 0.500 y: 0.158	
Accuracy, A	2.899	1.513	0.805	**1.739**

Table 2. Accuracy results.

When comparing the manual landmarking results, it can always be expected that there is a slight variation either due to user manual selection/mouse positioning and/or selection of landmarks [SSM06]. However, the pixel differences that was found between the automatic and manual landmarking are negligible.

5. CONCLUSION

In conclusion, we have successfully implemented an automatic landmarking method. The interactive tool is useful to visualise the execution of feature extraction and the localization process. The results have shown that after executing the method with

 ISBN 978-80-86943-71-8

different threshold pixel values, the identified output of the landmarks is similar. Therefore, we could label the landmark automatically.

Figure 7. Automatic landmarking on the tip of the nose

In the future, we will work on improving the detection and location of facial features by using a geometric model of the face and exploiting ratios to determine arbitrary landmarks or triangular features only on the face region. We will also evaluate the landmarks specifically in face registration. Finally, in the future, we will also conduct experiments on general object detection.

6. ACKNOWLEDGMENTS

This research is funded by Science Fund Malaysia grant number 01-01-09-SF0097. Thanks to UNIMAS and T.C. for the raw images. We also thank the reviewers for their valuable comments to improve the quality of the paper.

7. REFERENCES

[AABBO84] E.H. Adelson, C.H. Anderson, J.R. Bergen, P.J. Burt and J. M. Orgden. Pyramid methods in image processing. *In RCA Engineer, vol. 29(6).* pages 33-41. 1984.

[ANRS07] A. F. Abate, M. Nappi, D. Riccio and G. Sabatino. 2D and 3D face recognition: A survey. *Pattern recognition Letters*, vol.28, no. 14, pages 1885-1905, 2007.

[BJG85] P. J. Besl and R. C. Jain and Gupta. Three-dimensional object recognition. *ACM Computing Surveys (CSUR)*, no. 1, pages 75-145, 1985.

[BJ88] P. J. Besl and R. C. Jain. Segmentation through variable-order surface fitting. In Pattern Analysis and Machine Intelligence, IEEE Transactions on, 10(2), 167-192, 1988.

[CLDC12] H. H. Cui, W. H. Liao, N. Dai and X. S. Cheng. Registration and integration algorithm in structured light three-dimensional scanning based on scale-invariant feature matching of multi-source images. *Chinese Optics Letters*, no. 9, September 2012.

[Col06] D. Colbry. Human face verification by robust 3D surface alignment. *PhD thesis, Michigan State University*, 2006.

[CSJ05] D. Colbry, G. Stockman and A. Jain. Detection of anchor points for 3D face verification. In *Proceedings Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pages 118-118. IEEE, June 2005.

[CUS13] O. Çeliktutan, S. Ulukaya and B. Sankur. A comparative study of face landmarking techniques. *EURASIP Journal on Image and Video Processing*, 1:13-40, 2013.

[DSA08] H. Dibeklioglu, A. A. Salah and L. Akarun. 3D facial landmarking under expression, pose, and occlusion variations. In *Biometrics: Theory, Applications and Systems (BTAS 08), IEEE 2nd International Conference on*, pages 1-6. IEEE, September 2008.

[GKST10] P. Gupta, D. R. Kisku, J. K. Sing and M. Tistarelli. Maximized Posteriori Attributes Selection from Facial Salient Landmarks for Face Recognition. In *International Conference on Advances in Computer Science and Information Technology*, pages 337-344, 2010.

[Gor92] G. G. Gordon. Face recognition based on depth and curvature features. In *Proceedings of Computer Vision and Patt ern Recognition (CVPR 92), 1992 IEEE Computer Society Conference on*, pages 808-810. IEEE, June 1992.

[GZJ10] H. Guo, K. Zhang and Q. Jia. 2.5D SIFT Descriptor for Facial Feature Extraction. In Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 IEEE Sixth International Conference on, pages 723-726. IEEE, 2010.

[HYL08] Y. B. Han, J. Q. Yin and J. P. Li. Human Face Feature Extraction and Recognition Base on SIFT. *Computer Science and Computational Technology 2008 (ISCSCT08), International Symposium on*, vol. 1, pages 719-722. IEEE, 2008.

[JJC06] X. Ju, A. Jain and D. Colbry. Matching 2.5D face scans to 3D models. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(1):31-43, 2006.

[KD92] J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scales. *Image and vision computing*, no. 10(8), pages 557-564, 1992.

[KKK01] T. K. Kim, S. C. Kee and S. R. Kim. Real-time normalization and feature extraction of 3D face data using curvature characteristics. In *Proceedings of Robot and Human Interactive Communication, 2001 IEEE 10th International Workshop*, pages 74-79. IEEE 2001.

[Low04] D.G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 2004.

[MAVD03] A. B. Moreno, A. Sánchez, J. F. Vélez and F. J. Díaz. Face recognition using 3D surface-extracted descriptors. In *Irish Machine Vision and Image Processing Conference on*, vol. 2003. September 2003.

[NC09] P. Nair and A. Cavalloro. 3D face detection, landmark localisation, and registration using a point distribution model. In *Multimedia, IEEE Transaction on*, 11(4): 611-623, 2009.

[Ots75] N. Otsu. A threshold selection method from gray-level histograms. In *Automatica*, 11(285-296), 23-27, 1975.

[PPTTK09] P. Perakis, G. Passalis, T. Theoharis, G. Toderici and I. Kakadiaris. Partial Matching of interpose 3D facial data for face recognition. In Biometrics: Theory, Applications and Systems (BTAS09), IEEE 3rd International Conference on, pages 439-446. IEEE, 2009.

[PTK10] P. Perakis, G. Theoharis and I. A. Kakadiaris. 3D facial landmark detection and face registration. *Technical report, University of Athens, Greece*, 2010.

[PTPK09] P. Perakis, T. Theoharis, G. Passalis and I. Kakadiaris. Automaic 3D facial region retrieval fom multipose facial datasets. In *Proceedings 3D Object Retrieval, Eurographics Workshop on*, pages 37-44, April 2009.

[SAC09] P. Szeptycki, M. Ardabilian and L. Chen. A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking. In *Biometrics: Theory, Applications, and Systems (BTAS 09), IEEE 3rd International Conference on*, pages 1-6. IEEE 2009.

[SSM06] J. Shi, A. Samal, and D. Marx. How effective are landmarks and their geometry for face recognition? In *Computer Vision and Image Understanding,* 102:117–133, 2006.

[Sze12] P. Szeptycki. Processing and analysis of 2.5D face models for non-rigid mapping based face recognition using differential geometry tools. *PhD thesis, Ecole Centrale de Lyon*, 2012.

[Vau11] T. Vaughan. Multimedia: Making It Work. *McGraw – Hill Osborne Media, 8th ed*., page 142, 2011.

[TF95] E. Trucco and R. B. Fisher. Experiments in curvature-based segmentation of range data. In *Pattern Analysis and Machine Intelligence, IEEE Transactions* on, 17(2):177-182, 1995.

[TIC98] H. T. Tanaka, M. Ikeda and H. Chiaki. Curvature-based face surface recognition using spherical correlation, principal directions for curved object recognition. In *Proceedings of Automatic Face and Gesture Recognition, IEEE 3rd International Conference on*, pages 372-377. IEEE, April 1998.

[YT08] M. Yu and B. Tiddeman. Implementing the Scale Invariant Feature Transform (SIFT) Method. *http://www. cs. st-andrews. ac. uk/~ yumeng/yumeng-SIFTreport-5.18 _bpt. pdf*, 2008.

 ISBN 978-80-86943-71-8

Figure 8. The GUI of the automatic landmarking tool.

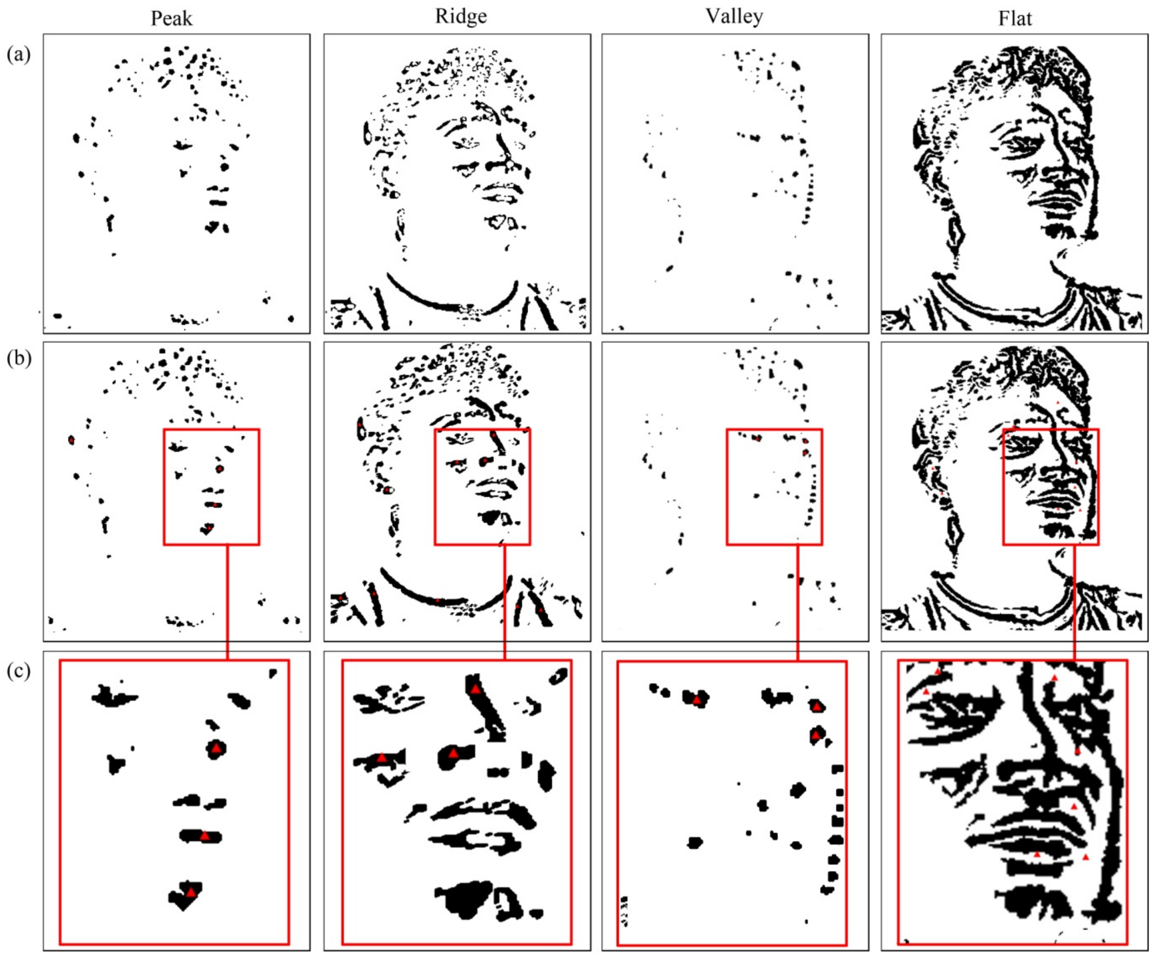

Figure 9. (a) Extracted features, (c) zoom-in/enlarged landmark areas of (b).

 ISBN 978-80-86943-71-8