Dense Facial Landmark Localization: database and annotation tool

1st Lei Ju
College of Automation
Southeast University
Nanjing, China
leiju1993@seu.edu.cn

2nd Xiangyu Zhu
National Laboratory of Pattern Recognition
Institute of Automation
Chinese Academy of Sciences
Beijing, China
xiangyu.zhu@nlpr.ia.ac.cn

3rd Zhen Lei
National Laboratory of Pattern Recognition
Institute of Automation
Chinese Academy of Sciences
Beijing,China
zlei@nlpr.ia.ac.cn

4th Xinfang Cui
College of Automation
Southeast University
Nanjing,China
1285550304@qq.com

5th Wankou yang
College of Automationy
Southeast Universit
Nanjing,China
wkyang@seu.edu.cn

6th Changyin sun
College of Automation
Southeast University
Nanjing,China
cysun@seu.edu.cn

***Abstract*—Databases are of great significance to researchers to achieve a satisfactory model. The lack of data is always a bottleneck to facial landmark localization, especially for the dense facial landmark detection. In this article, we provide a new dataset, called Dense Landmark Localization (DLL) database, which contains 39,198 images and is annotated in high quality. Annotating dense landmarks is a very tedious work due to two challenges. (a) Not every facial point has clear definition. Some of them distribute uniformly along the contour. Their labelled positions are determined by subjective judgement of the annotators, so that the quality of the annotation is poor. (b) Adjusting facial points one by one is time-consuming. The workload will increase dramatically when there are more points. To overcome the aforementioned problems, we propose a semi-automatic annotation tool to annotate dense points with much less clicks.**

***Index Terms*—landmark localizationdatabasesemi-automatic annotation tool**

I. Introduction

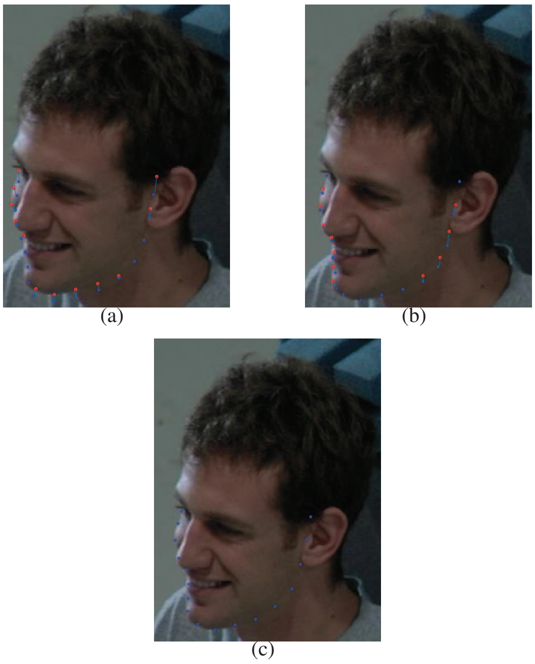
Facial landmark localization [1]–[3] has attracted a lot of researchers attention because of its extensive applications [4]–[7]. It has played a crucial part in face recognition, face attributes prediction, 3D face reconstruction, pose estimation and so on. It has also driven some interesting commercial applications, such as face morphing and virtual makeover.

Recently, great achievements and methodologies with good performance have been presented publicly. Many annotated facial landmark databases have also been provided by some researchers for free. These databases further promoted the development of the works about facial landmark localization. However, with the ringing and boom of deep convolutional neural networks and the higher requirements of face alignment methods on accuracy and robustness, the shortages of previous databases start to emerge. There are several concerns: (1) These open databases are often defined with different number of landmark points and semantic information. (2) The landmarks in some databases do not cover the whole face area, which can not be used for more complicated tasks. (3) Images in these databases are insufficient to learn millions of parameters of the neural networks in deep learning. In view of the above concerns, we take a significant step further and provide a new database with 39,198 images called Dense Landmark Localization (DLL) database, which is annotated with dense 84 facial points.

As we all know, manually annotating facial landmark points can be very expensive. The traditional way to make all alignments in one image needs fussy steps. Firstly, the annotator has to manually annotate enough images (if no existing database with the landmark configuration is available) to train a initial model. After that, the annotator get the landmark localization by using the model which are achieved in the first step. After getting the rough position of landmarks, the annotator clicks and drag every landmark to the right position which depends on his/her subjective judgement. Hence most landmarks do not have explicit semantic definition and they should be evenly distributed between two semantic landmarks, the annotator always needs to adjust the coordinates of some landmarks to make them look evenly spaced just as Fig. 1 shows. Finally, the annotator submits the annotation and switches to the next image.

The annotation process above is easy but time-consuming. It involves clicking on the original landmark, dragging the landmark to imaginary places and adjusting these points to make sure they are evenly distributed. All of these steps require annotators to work with high concentration and the labelled positions are largely dependent on annotators own imagination. Actually annotators have to carry out the above progress repetitively at most time, because it is almost impossible to achieve a perfect annotation result after the first modification (especially these points with uniform distribution). On the other hand, long time accurate clicks and mental imaginary

978-1-7281-3936-4/19/$31.00 ©2019 IEEE



(a) (b) (c)

Fig. 1. The traditional way to annotate an image. (a) The initial points are shown in blue, these red points are the right positions an annotator may choose. (b) To make these points spread uniformly along the contour, an annotator needs adjusting them again. This step can be repeated timely. (c) The final result

make the annotators enter fatigue rapidly. S. M. Kosslyn and R. N. Shepard have demonstrated that mental imagery has a cognitive cost in [8], [9]. The mental fatigue will dramatically increase annotators reaction time. In this paper, we propose a semi-automatic methodology for facial landmark annotations. This approach makes full use of the facial structures as the priori, which can reduce the cognitive cost of annotators. We will release this semi-automatic annotation tool and part of DLL database publicly.

II. Related Work

In this section we discuss related works in facial databases and semi-automatic methodologies for facial landmark annotations.

Public facial databases. Facial landmark localization has been a very popular topic over the past years. The performance of the trained models are stuck due to the unavailable of large high quality databases. Many groups have made significant contributions to overcome this problem by providing more databases. S.Milborrow et al. published the MUCT data set containing 3,755 images with 76 annotated points, this database offers more diversity of lighting, age, and ethnicity [10]. In [11], the CMU Multi-Pose, Illumination and Expression (Multi-PIE) Database contains around 750,000 images of 337 subjects captured under laboratory conditions, the accompanying facial landmark annotations consist of a set of 68 points. In 2016, Ankan Bansal et al. released a data set, called UMDFaces, which has 367,920 face annotations of 8501 subjects [12]. The author used Amazon Mechanical Turk (a widely used crowd-sourcing platform to get human annotations) to provide 21 annotations for these images. Besides the datasets above, there are many other good databases, such as XM2VTS [13], AR [14], LFPW [15], HELEN [16], AFW [17] and IBUG [1], just to name a few. The images these databases contained and the number of annotations they provided are shown in Table I.

TABLE I
THE EXISTING DATABASES PROVIDED BY FORMER RESEARCHERS

Name of Database	# of Images	# of Points
MUCT	3755	76
XM2VTS	2,360	68
AR	4,000	22
LFPW	1,035	35
HELEN	2,330	194
AFW	468	6
IBUG	135	68
UMDFaces	367,920	21
DLL	**39,198**	**84**

Semi-automatic methodologies for facial landmark annotations. In 2012, Yan Tong et al. proposed an approach estimating the locations of a set of landmarks for a large image ensemble [18]. A shape model is learned online to constrain the landmark configuration. However, the limitation of this methodology is obvious, this tool has only been applied on images that are captured under controlled conditions. In [19], Christos Sagonas et al. solved this problem with the help of Active Orientation Models (AOMs). In [20], the author used this annotation tool to annotate Menpo database by replacing AOM with Mnemonic Descent Method(MDM) [21].

Finally, our work was inspired by the recent paper by Dim P. Papadopoulos et al. [22] on drawing a bounding box by clicking on several extreme points. The annotation time is much less than drawing boxes with traditional way . Based on this idea, we propose a semi-automatic methodology for facial landmark annotations that requires annotators to click several points (far less than the total of facial landmark points) on the outline of the face.

III. Dense Landmark Localization Database

In this section, we first explain the definition of our proposed 84 points and demonstrate the superiority of our definition by comparing with definition of 68 points. Then we describe the images collection procedure and statistics of DLL Database. Finally, we set a evaluation protocol for the proposed database.

A. Points Definition

Before presenting our definition, we elaborate the definition of the widely used 68 points. As Fig. 2(a) shows, the 68 points include 17 points for face contour, 10 points for eyebrows, 12 points for eyes, 9 points for nose and 20 points for mouth. Intuitively, this definition gives no description points on lower part of eyebrows and nosewing. Besides that, 6 points are slightly thin to represent an eye contour.

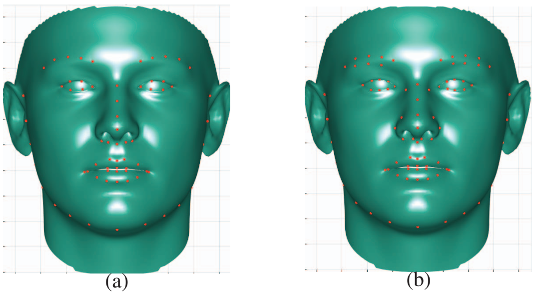


Fig. 2. The definition of 68 points(left) and 84 points(right)

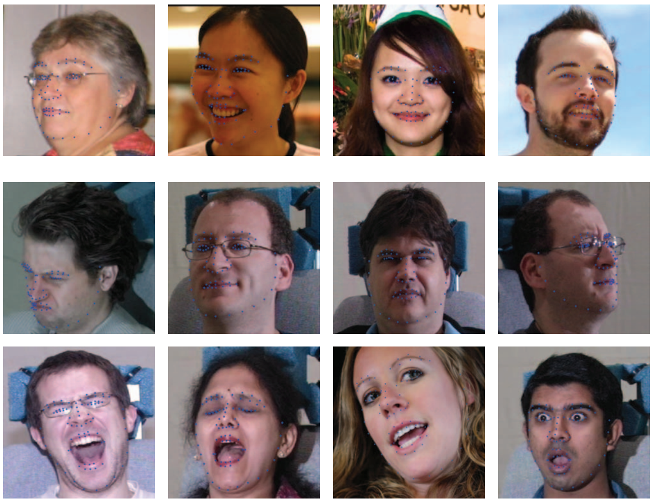


Fig. 3. The images and points we provided in DLL database, the first and second rows show variations in yaw angles and pitch angles, the last row describe the extreme expressions provided in DLL database

To alleviate these limitations, we propose 84 points whose definition are shown in Fig. 2(b). In our definition, 6 more points are used to make up the excalation of eyebrows and 4 more points to describe eyes. We also added the definition of nosewing with 6 points and fine-tune the definition of nose to make them have more explicit semantics.

B. Database Composition

To construct DLL database, we take 3,699 images from 300-W challenge, 15,924 images form Multi-PIE and 135 images from IBUG, we also collected 19,440 images from web to strengthen the variation in pose, illumination, expression and background. In totally, 39,198 images are available in DLL database.

We gathered these images together and re-annotated them with the help of our proposed semi-automatic annotation tool (section 4). Fig. 3 shows some samples from DLL database with 84 annotated points. For those faces with yaw angles, the defined points may be unseen in 2D plane, we annotated these points on the face silhouette instead.

At last, we propose a evaluation protocol on DLL database. We set 34,398 images for training and 4,800 images for testing, each subject is divided into three parts: Left, Front and Right, a face with its yaw angle ranging in $[-15°, 15°]$ is identified

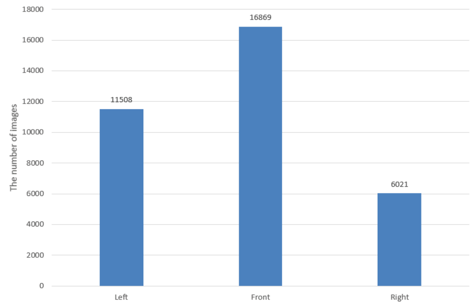


Fig. 4. The number of images in each subject

TABLE II
THE VARIATIONS OF PITCH AND YAW ANGLES IN DLL DATABASE

pose	$-30° : -15°$	$-15° : -0°$	$0° : 15°$	$15° : 30°$
Pitch	14.59%	61.05%	24.02%	0.34%
Yaw	19.11%	26.02%	22.00%	32.87%

as a front face, the yaw angle bigger than $15°$ or smaller than $-15°$ is classified as Left or Right, respectively. The pose distribution of DLL is shown in Fig. 4 and Table II. Besides, DLL has rich variations in expression, including many images with closed eyes, largely opened mouth and, which is shown in Table III.

IV. SEMI-AUTOMATIC ANNOTATION TOOL

In this section, we propose a semi-automatic dense landmark annotator based on nonrigid ICP [23] and 3DMM [24], which provides a faster and more accurate way to annotate the large database.

A. 3D Morphable Model

3D Morphable Model (3DMM) [25] is proposed to describe faces in 3D space. Chu et al. [26] extended 3DMM to contain expression as the offset to the neutral face. By training on the 3D face scans with neutral expression, the principal axis of variations can be gotten and denoted as A_{id}. By training a PCA [27] on the offsets between expressive scans and neutral scans, the author get another principal axis A_{exp}. The neutral faces S can be shown as:

$$S = \overline{S} + A_{id}\alpha_{id} + A_{exp}\alpha_{exp}. \tag{1}$$

Where the $\overline{S}$ is the mean shape, α_{id} and α_{exp} are the shape weight and expression weight respectively. In [28], xiangyu zhu et al. combined the 3DMM with facial key points by

TABLE III
THE PERCENTAGE OF SPECIFIC EXPRESSIONS IN DLL DATABASE

Pose	Open Eyes	Closed Eyes	Open Mouths	Closed Mouth
Left	74.01%	25.99%	45.18%	54.82%
Front	89.55%	10.45%	42.12%	57.88%
Right	72.64%	27.36%	43.43%	56.57%

using the weak perspective projection. The author got the identity shape A_{id} from the Basic Face Model (BFM) [29] and employed the expression A_{exp} from the Face Warehouse [30]. The author deformed the face model onto the image plane by solving

$$s_{2d} = fPR\left(\alpha, \beta, \gamma\right)\left(S + t_{3d}\right). \tag{2}$$

In the above equation, s_{2d} is facial key points position, f is the scale factor, P is the orthographic projection matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, R is the rotation matrix. Assuming having the corresponding 3D facial landmark on the face model, the 2D coordinates of 3D points are denoted as s_{2dt}. Naturally, the distance between s_{2d} and s_{2dt} should be small. So the author minimized

$$arg \min_{f,R,t_{3d},\alpha_{id},\alpha_{exp}} \left\| s_{2dt} - s_{2d} \right\|. \tag{3}$$

The author did this on the Multi-PIE database and had achieved a robust and good performance.

B. Nonrigid ICP for Curve Adjustment

Nonrigid ICP is an extension of ICP framework. Brian Amberg et al. [23] retained the convergence properties of the original algorithm and applied it to nonrigid registration in 3D-space. In this paper, we modify nonrigid ICP to perform the registration between points and the target facial curve. Assume we have the target curve V and s set of initial facial points $P = [p_1, p_2, ..., p_n]$the line between neighbouring points denote as ϵ. Our goal is to find parameters X to adjust these points to the target curve and keep the correlation between them at the same time. The cost function of this progress contains three parts: distance term, stiffness term and simple landmark term. The full cost function is defined as

$$E\left(X\right) := E_d\left(X\right) + \alpha E_s\left(X\right) + \beta E_l\left(X\right). \tag{4}$$

The parametrization of the mapping for every facial key points is X_i, which is one affine 2x3 transformation matrix. The X denotes the unknown parameters and $X := \begin{bmatrix} X_1 & ... & X_n \end{bmatrix}^T$. Normally, to make the facial key points close to the target curve, the distance between the initial points and target surface should be small. The first term of the cost function E_d limits this and is represented as

$$E_d\left(X\right) := \sum_{1}^{n} w_i dist^2\left(V, X_i p_i\right). \tag{5}$$

Where the $p_i = \begin{bmatrix} x & y & 1 \end{bmatrix}^T$. The $dist\left(V, p\right)$ represent the distance between a initial point p and its closest point on the target curve. If no corresponding point is available, the w_i is set to zero. Denoting the correspondences as (p_i, u_i), then the equation (5) becomes

$$\bar{E}_d\left(X\right) := \sum_{1}^{n} w_i \left\| X_i p_i - u_i \right\|^2 \tag{6}$$

$$= \left\| \left(W \bigotimes I_3\right) \left(\begin{bmatrix} X_1 & & \\ & ... & \\ & & X_n \end{bmatrix} \begin{bmatrix} p_1 \\ ... \\ p_n \end{bmatrix} - \begin{bmatrix} u_1 \\ ... \\ u_n \end{bmatrix} \right) \right\|^2. \tag{7}$$

In the above equation, $W := diag\left(w_1, ..., W_n\right)$ is a $n \times n$ identity matrix. Define sparse matrixes D and U as

$$D = \begin{bmatrix} p_1^T & & & \\ & p_2^T & & \\ & & ... & \\ & & & p_n^T \end{bmatrix}, \tag{8}$$

$$U := \begin{bmatrix} u_1 & ... & u_n \end{bmatrix}^T. \tag{9}$$

Then the distance term can be written as

$$\bar{E}_d\left(X\right) = \left\| W\left(DX - U\right) \right\|_F^2. \tag{10}$$

The second term E_s is stiffness term, which penalises the weighted difference of neighbouring facial points with the Frobenius norm and a weighting matrix $G := diag\left(1, 1, \gamma\right)$ to keep the correlation during the transformation. It is denoted as

$$E_s\left(X\right) := \sum_{\{i,j\} \in \epsilon} \left\| \left(X_i - X_j\right) G \right\|_F^2. \tag{11}$$

In our progress, the γ is set to 1. To express the relationship between the neighbouring vertices, we use the node-arc incidence matrix M to describe the topology just as [23] did. Then the stiffness term can be written as

$$E_s\left(X\right) = \left\| \left(M \bigotimes G\right) X \right\|_F^2. \tag{12}$$

Assuming we have a set of anchor points $A := \{(p_1, a_1), ..., (p_l, a_l)\}$ which can be given by annotators, the last cost term is defined as

$$E_l\left(X\right) := \sum_{i \in anchor} \left\| X_i p_i - a_i \right\|^2. \tag{13}$$

Taking the corresponding rows out of D to organize these rows in D_A and making $U_A = \begin{bmatrix} a_1 & ... & a_l \end{bmatrix}^T$, then similar to the first term, the landmark term can be

$$E_l\left(X\right) = \left\| D_A X - U_A \right\|_F^2. \tag{14}$$

The whole cost function will be :

$$\bar{E}\left(X\right) = \left\| \begin{bmatrix} \alpha M \bigotimes G \\ WD \\ \beta D_A \end{bmatrix} X - \begin{bmatrix} 0 \\ WD \\ D_A \end{bmatrix} \right\|_F^2 \tag{15}$$

$$= \left\| AX - B \right\|_F^2. \tag{16}$$

The function above can be minimized directly and exactly. Nonrigid ICP has great robustness to initial conditions and missing data and the optimal progress is very effective.

Fig. 5. The calibration points for aligning a rough position

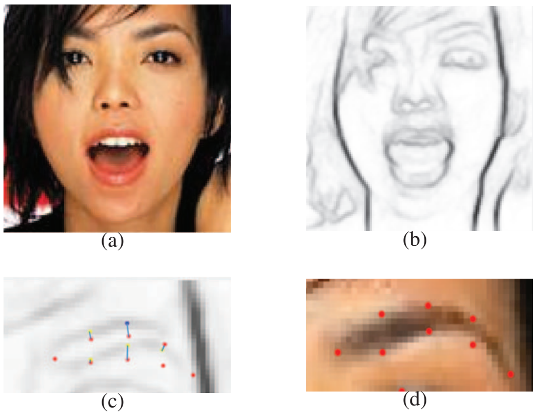


Fig. 6. Annotating the eyebrow. (a) The original image which was scaled to 120×120. (b) The edge detected by the detection method. (c) red points are the initial points, the blue point is the directing point which was given by annotator and these yellow points are the closest points searched on the edge within 2 pixels. The stiffness term and simple landmark term can regularize these points with wrong closest points or with no corresponding points. (d) The final result

C. Proposed Tool

The main idea of our semi-automatic tool is to employ a strong edge detection method to get the facial edge information and then we adjust the initial facial key points to the edge with the help of nonrigid ICP. After the facial features have been corrected, we use these points to get a more accurate initialization for the facial counter by 3DMM.

We divide the tool into two parts, one part is for the facial features and the other one is for the facial counter. Firstly, for the facial features part, before adjusting the initial points, we use a fast and robust edge detection model proposed by piotr Dollar et al. [31] to get the edge information in the image. Then the annotator will be asked to click on some calibration points on the face, these anchor points can make sure that the final annotations will not be far away from their corresponding facial parts. Besides, they reduce the number of wrong edge searching points when we perform nonrigid ICP. We define 15 calibration points (distributed on eye-brow tips, eyes corners, mouth corners, the chin point and the upper contact points between the ears and the face, which are shown in Fig. 5). After that, the annotator can click on any places to provide more anchor points t guide the corresponding initial points, the rest points will be adjusted automatically by nonrigid ICP algorithm.

Taking eyebrows as an example, the annotators need to click on the tip points of the eyebrows, then the tool adjusts the points with non-reflective similarity transformation to make sure that the initial points are near the eyebrows. After that, a directive point which is corresponding to one of the initial eyebrow points should be given to perform an nonrigid ICP. The other points will search for an corresponding edge points within 2 pixels (see Fig. 6). The stiffness term in NICP can avoid the wrong edge points misleading the adjusting direction. This progress can be repeated until a good annotation is done.

The other parts on faces are annotated as the eyebrows. Note that we do not design a tool for the nose part, because the nose part has clear semantic information but containing few contour information. This means that the nose points are relatively easy for a trained model to initialize. In this article, we fine-tune the nose points with the traditional way.

The second part is the face contour which has 17 points in our definition. The traditional annotators have to spend a lot of time on labelling these points since they do not have clear semantic positions. In our semi-annotation tool, we employ 3DMM to assist the annotation. Similar to the first part, we also set several specific points for annotators to click on, they are the chin point and the upper contact points between the ears and the face. The 3DMM take these three points and the labelled facial feature points as the input, then it fits the 3D face model and provides fine-grained projected 3D contour points, which is shown in Fig. 7 shows. However, these points can not provide shape information of contour line, so we perform edge detection and nonrigid ICP again to constrain these points to the right line.

D. Efficiency

The tool is realized with the GUI in MATLAB 2015b and can be very efficiency. Before we test its performance, we cut and scale these images to 120×120 pixels to reduce edge detection time by using the face detection box provided by fastrpn.

We use a subset of UMDFaces whose landmarks have been roughly initialized. The annotator takes 43.5 minutes to annotate 50 images, more specifically, the annotator spends about 1630 seconds on annotating facial features except nose, 450 seconds to adjust nose and 530 seconds on annotating face contour. With the traditional way, an expert annotator needs around 5 min to annotate 68 points on one face. While with our tool, an annotator is able to annotate 84 points within 1 minute which is about 5 times faster than the traditional way.

V. Conclusion

In this paper, we propose 84 landmarks definition which is more dense and comprehensive than the traditional 68 landmarks. Besides, we provide a new facial landmark database

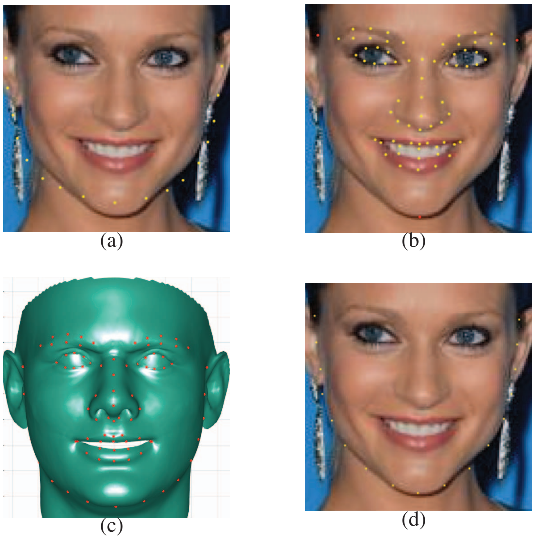


Fig. 7. Re-initialize the contour points. (a) Original initial points. (b) Input of 3DMM, the three red points are provided by annotator and the rest are achieved from the first part. (c) The 3D model calculated by 3DMM. (d) New initialization predicted by the output of 3DMM

DLL which consists of 39,198 images from Multi-PIE with rich variations in pose and expression. The DLL is much larger than existing dense landmark datasets and can help to train better face alignment methods. In addition, we present a semi-automatic annotation tool to help annotators. Based on nonrigid ICP and 3D Morphable Model, the toll is able to help annotators annotate 84 points with much few clicks. Finally, we compared our tool with the tradition way, tradition way needs 5 mins to annotate 68 points. While our semi-annotation tool reduces this cost to 1 min to annotate 84 points with no damage on quality.

VI. Acknowledgments

This work was supported by the Chinese National Natural Science Foundation Projects #61806196, #61876178.

References

[1] Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., Pantic, M.: 300 faces in-the-wild challenge: The first facial landmark localization challenge. In: IEEE International Conference on Computer Vision Workshops. (2013) 397–403
[2] Sun, Y., Wang, X., Tang, X.: Deep convolutional network cascade for facial point detection. In: IEEE Conference on Computer Vision and Pattern Recognition. (2013) 3476–3483
[3] Lai, H., Xiao, S., Cui, Z., Pan, Y., Xu, C., Yan, S.: Deep cascaded regression for face alignment. Computer Science (2015)
[4] Chung, J.S., Zisserman, A.: Lip Reading in the Wild. (2017)
[5] Booth, J., Roussos, A., Zafeiriou, S., Ponniahy, A., Dunaway, D.: A 3d morphable model learnt from 10,000 faces. In: IEEE Conference on Computer Vision and Pattern Recognition. (2016) 5543–5552
[6] Booth, J., Roussos, A., Ponniah, A., Dunaway, D., Zafeiriou, S.: Large scale 3d morphable models. International Journal of Computer Vision (5) (2017) 1–22
[7] Shu, Z., Yumer, E., Hadap, S., Sunkavalli, K., Shechtman, E., Samaras, D.: Neural face editing with intrinsic image disentangling. In: IEEE Conference on Computer Vision and Pattern Recognition. (2017) 5444–5453
[8] Kosslyn, S.M., Thompson, W.L., Klm, I.J., Alpert, N.M.: Topographical representations of mental images in primary visual cortex. Nature **378**(6556) (1995) 496–498
[9] Shepard, R.N., Metzler, J.: Mental rotation of three-dimensional objects. Science **171**(3972) (1988) 598–599
[10] Milborrow, S., Morkel, J., Nicolls, F.: The muct landmarked face database. (2010)
[11] Gross, R., Matthews, I., Cohn, J., Kanade, T.: Multi-pie. In: IEEE International Conference on Automatic Face & Gesture Recognition. (2008) 1–8
[12] Bansal, A., Nanduri, A., Castillo, C., Ranjan, R., Chellappa, R.: Umd-faces: An annotated face dataset for training deep networks. (2016)
[13] Messer, K.: Xm2vtsdb: The extended m2vts database. Proc.intl.conf.on Audio & Video Based Biometric Person Authentication (1999) 72–77
[14] Martinez, A.M.: The ar face database. Cvc Technical Report **24** (1998)
[15] Belhumeur, P.N., Jacobs, D.W., Kriegman, D.J., Kumar, N.: Localizing parts of faces using a consensus of exemplars. IEEE Transactions on Pattern Analysis & Machine Intelligence **35**(12) (2013) 2930–2940
[16] Le, V., Brandt, J., Lin, Z., Bourdev, L., Huang, T.S.: Interactive facial feature localization. In: European Conference on Computer Vision. (2012) 679–692
[17] K?stinger, M., Wohlhart, P., Roth, P.M., Bischof, H.: Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In: IEEE International Conference on Computer Vision Workshops. (2012) 2144–2151
[18] Tong, Y., Liu, X., Wheeler, F.W., Tu, P.H.: Semi-supervised facial landmark annotation . Computer Vision & Image Understanding **116**(8) (2012) 922–935
[19] Sagonas, C., Antonakos, E., Tzimiropoulos, G., Zafeiriou, S., Pantic, M.: 300 faces in-the-wild challenge: database and results * **. Image & Vision Computing **47** (2016) 3–18
[20] Zafeiriou, S., Trigeorgis, G., Chrysos, G., Deng, J., Shen, J.: The menpo facial landmark localisation challenge: A step towards the solution. In: Computer Vision and Pattern Recognition Workshops. (2017) 2116–2125
[21] Trigeorgis, G., Snape, P., Nicolaou, M.A., Antonakos, E., Zafeiriou, S.: Mnemonic descent method: A recurrent process applied for end-to-end face alignment. In: Computer Vision and Pattern Recognition. (2016)
[22] Papadopoulos, D.P., Uijlings, J.R.R., Keller, F., Ferrari, V.: Extreme clicking for efficient object annotation. In: IEEE International Conference on Computer Vision. (2017) 4940–4949
[23] Amberg, B., Romdhani, S., Vetter, T.: Optimal step nonrigid icp algorithms for surface registration. In: Computer Vision and Pattern Recognition IEEE Computer Society Conference on. (2007) 1–8
[24] Blanz, V., Vetter, T.: Face recognition based on fitting a 3d morphable model. IEEE Transactions on Pattern Analysis & Machine Intelligence **25**(9) (2003) 1063–1074
[25] Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d faces. In: Conference on Computer Graphics and Interactive Techniques. (1999) 187–194
[26] Chu, B., Romdhani, S., Chen, L.: 3d-aided face recognition robust to expression and pose variations. In: IEEE Conference on Computer Vision and Pattern Recognition. (2014) 1907–1914
[27] Wold, S., Esbensen, K., Geladi, P.: Principal component analysis. Chemometrics & Intelligent Laboratory Systems **2**(1) (1987) 37–52
[28] Zhu, X., Lei, Z., Yan, J., Yi, D., Li, S.Z.: High-fidelity pose and expression normalization for face recognition in the wild. (2015) 787–796
[29] Paysan, P., Knothe, R., Amberg, B., Romdhani, S., Vetter, T.: A 3d face model for pose and illumination invariant face recognition. In: IEEE International Conference on Advanced Video and Signal Based Surveillance. (2009) 296–301
[30] Cao, C., Weng, Y., Zhou, S., Tong, Y., Zhou, K.: Facewarehouse: A 3d facial expression database for visual computing. IEEE Transactions on Visualization & Computer Graphics **20**(3) (2013) 1
[31] Dollr, P., Zitnick, C.L.: Structured forests for fast edge detection. In: IEEE International Conference on Computer Vision. (2014) 1841–1848