メッシュモデルを用いた顔面非対称性の定量評価

プジタ　グナラトネ　　佐藤幸男

名古屋工業大学　電気情報工学科
〒466-8555　名古屋市昭和区御器所町

E-mail: pujitha@Hilbert.elcom.nitech.ac.jp, sato@elcom.nitech.ac.jp

あらまし：本稿では、キュービック・フェイサーにより様々な表情の距離画像
とテクスチャ画像を計測し、カラー画像から抽出された顔面の特長点と距離画
像より得られる 3D データにより標準顔メッシュモデルで表現する手法につい
て述べる。この、標準顔メッシュモデルへの当てはめ過程では、多項式モデル
を用いた最小二乗法による補間を行っている。また、標準顔メッシュモデルの
左右での非対称性を定量的に評価する手法を提案する。本研究では、これらの
標準顔メッシュモデルの左右での非対称性を用いることにより顔面神経麻痺の
病状の評価を目指す。

Quantification of Facial Asymmetry with Mesh based Approach.

Pujitha Gunaratne　　　　Yukio Sato

Department of Electrical and Computer Engineering
Nagoya Institute of Technology
Gokiso, Showa, Nagoya 466-8555, Japan.

E-mail: pujitha@hilbert.elcom.nitech.ac.jp, sato@elcom.nitech.ac.jp

Abstract: In this work, we present a method of quantifying facial asymmetry using a
topological face mesh, which is adapted to a measured 3D data, acquired by the
Cubicfacer rangefinder system. Pre-determined facial actions are measured and the
generic face mesh is adapted to the measured 3D data by extracting feature points in
corresponding color image. The mesh adaptation process approximates a polynomial
curve using least-squares estimation. Variances of identical patches on left and right
sides of the face are analyzed with normal and disordered human subjects to generate an
estimation of facial asymmetry. Quantification measure is then computed against a
variance distribution of normal subjects.

1. Introduction

Although the early modeling process of human face largely dealt with animating head models to generate CG characters [1][2], with the advent of techniques to generate 3D models and accurate acquisition of 3D data by rangefinder systems, the application areas of human facial modeling have spanned rapidly. Present day work on facial analysis is directly applied in numerous areas ranging from face recognition, morphing, character simulation and animation, and medical applications. In this work our attention is primarily focused on analyzing facial expressions of human subjects to determine the asymmetry presented in each individual. The estimation is expected to use in clinical analysis of facial nerve paralysis treatments. Some attempts have been made to estimate facial asymmetry based on video captured sequences of facial expressions by employing 2D image processing techniques [3]. Thus they inherently loose true 3D shape information of facial deformation in expressions.

Although in practice it is hard to determine the faces with ideal symmetry, we consider this property as *apparent-symmetry*, interpreted by human observers. Our system measures individuals of normal as well as expression disordered with different facial actions and produce 3D as well as color texture images. Each measured set of 3D range data is then adapted to a pre-designed generic face mesh model. This process involves the extraction of determined feature points from the color image and poling those to the identified nodes on the face mesh. Then the nodal points representing non-feature points are mapped with interpolating a polynomial function approximated by a least squares estimator. This procedure is applied to the each facial action measured, thus producing a complete wrap of the generic face mesh on the measured data. Facial deformation analysis is based on the adapted mesh patches. Left and right side patches are compared for their deformation and a patch variance is calculated for every patch.

There are various approaches of generic face mesh deformation are proposed in the past that applied in parameterized and control point models [4,5,6], spline based model [7] etc. In most cases, mesh adaptation require segmentation of underlying 3D surface or setting up control points on feature boundaries, generating overheads in processing.

In quantifying facial asymmetry, we need to have a control set to compare different facial actions. Therefore we first measure six subjects with apparent expression symmetry and take that as the control set of measurement. Preceding measurements are compared with respect to this control set to determine the presence of asymmetry. The regression analysis is performed on the data sets and the standard error factor of the control set is calculated. This factor is used to quantify the degree of asymmetry in other faces.

2. The Approach

First we measure each individual with pre-determined facial actions using the Cubicfacer rangefinder system, which consists of two laser range scanners mounted laterally with respect to the face, and a CCD camera facing the measuring profile (fig.1). The face is scanned complementary by the two laser scanners and the complete frontal face model is generated within a second [8]. The adaptive mesh generation followed by texture mapping produces a realistic 3D model of the frontal face as in fig.2. The characteristics of this measuring geometry, where a single camera with identical orientation is used for the measurement of both range and texture images makes it 1:1 correspond to each other.
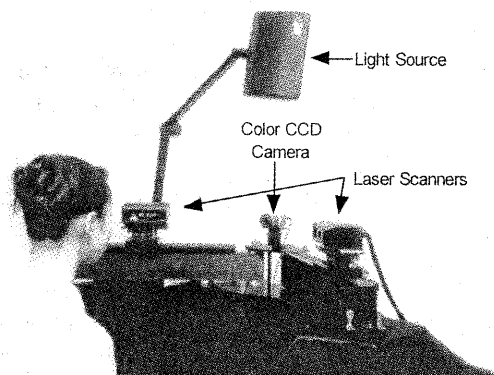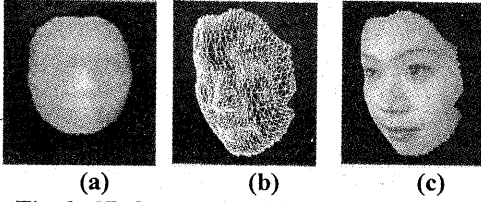

Fig1: Face measurement system

Fig. 2: 3D face model. (a) Median filtered range image (b) Adaptive mesh result (c) Texture mapped result

Fig.4: Points selected for mesh adaptation. (a) Texture points (b) Identical mesh nodes.

During the texture mapping, adaptive mesh is generated on the range data. Since the adaptive mesh generated does not possess consistent triangle density on both sides of the face (fig.2(b)), it is hardly suitable for interpretations based on symmetry features of the face. Therefore, we adopt an arbitrary generic mesh (fig.3) that is symmetric along the median plane, which is the vertical plane passes through the center of the nose, cuts the face into identical left and right halves. This generic mesh is in the 2D from, lies on the XY plane. Therefore we adopt a method of wrapping the mesh on to the measured 3D range data with the use of the corresponding color texture image.
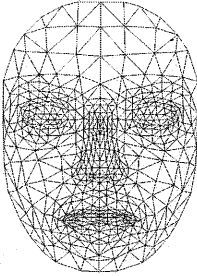

Fig.3: Generic face mesh

2.1 Mesh adaptation

The mesh adaptation can be time consuming, tedious process if it involves segmentation of range data to extract features. Instead, here we apply a simple method of extracting features by using the corresponding color image, since it possesses the property of 1:1 correspondence with the range image. We select 42 pre-determined points manually, on the color image, which correspond to mapping points on the face mesh (fig.4), for the adaptation process.
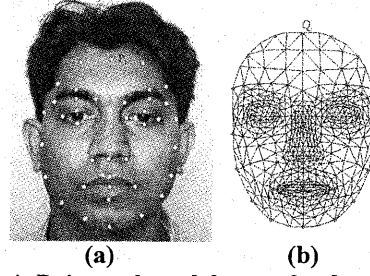
Points extracted from the color image are then mapped with mesh nodes by approximating a polynomial function using a least squares estimator.

2.2 Least squares approximation

An Nth order polynomial function is approximated to the range data using a least squares estimator to adapt the generic mesh to the measured data. This process consists of two steps. First we move mesh vertices to the extracted points of the color image, and then poll the Z values from the corresponding range image.

Let us consider the parametric function given by,

$$Z = f(x, y).$$

Where, $f(x, y)$ represents by a polynomial of Nth degree, given by,

$$f(x, y) = a_{00} + \sum_{j=1}^{N} \sum_{i=0}^{j} a_{j-i,i} x^i y^{j-i} \text{------(1)}$$

When $N=2$, it takes the form,

$$f(x, y) = a_{00} + \sum_{i=0}^{1} a_{1-i,i} x^i y^{1-i} + \sum_{i=0}^{2} a_{2-i,i} x^i y^{2-i}$$

$$\text{----------------------------(2)}$$

Now consider match points Pi and Qi, where $i=1....n$, represent points on the color image and the mesh respectively (Fig.4). Pi's are extracted from the color image and Qi's are known with respect to the topological mesh. Let (x_{p_i}, y_{p_i}) and (x_{Q_i}, y_{Q_i}) represent 2D coordinates of Pi and Qi respectively.

We can thus calculate the displacement vectors, $\mathbf{dx}_i = (x_{P_i} - x_{Q_i})\vec{\mathbf{m}}$ and $\mathbf{dy}_i = (y_{P_i} - y_{Q_i})\vec{\mathbf{n}}$, where $\vec{\mathbf{m}}$ and $\vec{\mathbf{n}}$ are unit vectors along x and y direction respectively, for all matching points $i = 1, \ldots, n$.

Since we extract 42 points for initial matching, n is set to 42. To calculate displacement vectors for the entire data set of the mesh, we approximate the parametric function given in eq.(1) using the least squares method, polling $\mathbf{dx}_i$ and $\mathbf{dy}_i$ in Z axis as depicted in fig.5.
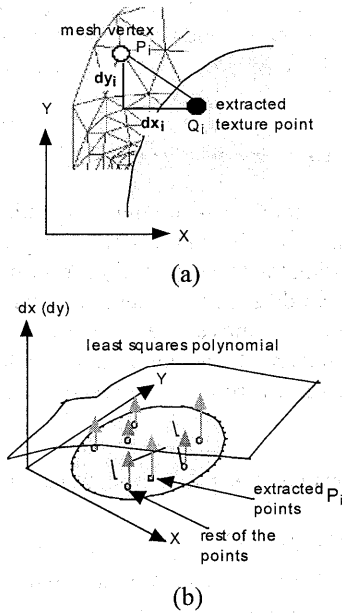

(a)

(b)

Fig.5. Least squares polynomial approximation.
(a) Displacement vector calculation.
(b) Polynomial approximation.

Thus, coefficients a_{00}, a_{10}, a_{01}, a_{20}, a_{11}, a_{02} can be calculated using displacement vectors and therefore the nodal displacements of other points can be calculated by simply interpolating the polynomial function.

We repeat this procedure again by increasing the order of the polynomial to move the mesh points further closer to the expected locations by iterative approximation. We then separate feature points on different regions of the face, namely eye, nose and mouth regions, where a high concentration of facial features is observed.

This local matching is done to ensure a better mapping for the prominent feature areas of the face. Finally, Z values are mapped from the corresponding range values, since both color and range images have 1:1 correspondence, producing a complete wrap of the face mesh to the measured 3D data.

Once the 3D mesh is generated, we apply asymmetry measurements against the measured facial actions to estimate the difference of deformation on both sides of the face.

3. Estimation of Facial deformation

Facial deformation is estimated by calculating variances of patches of the adapted mesh, on both sides of the face.

Patch variances are estimated with respect to the sub-meshes representing different regions of the face. Forehead, eye, nose and mouth meshes are defined in the generic face mesh beforehand, and used to estimate the patch variances. Consider two match patch pairs in a given sub-mesh marked as P_{Li} and P_{Ri}, representing left and right side patches respectively (fig.6). Their corresponding edge lengths are denoted as ξ_{Li} and ξ_{Ri} respectively, where $i = 1, 2, 3$.
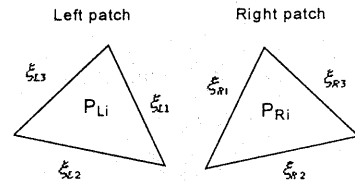

Fig. 6: Matching patch pairs.

If the variance of the ith patch is σ_i^2 and total path variance of a given sub-mesh with N patches is σ^2, we can define,

$$\sigma^2 = \sum_{i=1}^{N} \sigma_i^2 = \sum_{i=1}^{N} \sum_{j=1}^{3} \left\| \xi_{Li_j} - \xi_{Ri_j} \right\|^2.$$

Thus, we measure the variance in terms of sub-mesh patches.

In the applications of asymmetric facial expression analysis, it is often required to measure the relative variances of different facial actions and compare them (Fig.7).
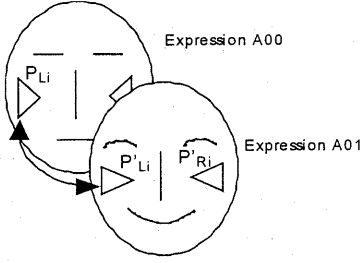
Fig.7. Variance comparison of two expressions.

In a similar calculation, as done in the previous case, suppose patch P_{Li} of expression A00 occupies the patch P'_{Li} in expression A01. Let the patch variances of left and right sides denote $\sigma_L{}^2$ and $\sigma_R{}^2$ respectively. Thus,

$$\sigma_L{}^2 = \sum_{i=1}^{N_L}\sum_{j=1}^{3}\left\|\xi_{Li_j} - \xi_{L'i_j}\right\|^2 \quad \text{and}$$

$$\sigma_R{}^2 = \sum_{i=1}^{N_R}\sum_{j=1}^{3}\left\|\xi_{Ri_j} - \xi_{R'i_j}\right\|^2 \quad \text{where } \xi_L, \xi_{L'}, \xi_R,$$

$\xi_{R'}$ represent the lengths of the same patch in right and left sides in different expressions. N_L and N_R represent the number of patches in left and right sides of the same sub-mesh. Then the comparison is performed for both sides of the face to detect the asymmetry.

4. Quantitative analysis

In the experiment we measure two facial actions, namely eye closure and grin, of different human subjects with no apparent expression disorders, as well as subjects with some expression disorders. These actions are chosen to cover movements of most parts of the face. We analyze the range data of human subjects measured by the cubicfacer range finder system, using above described variance estimations.

The subjects are first measured at the relaxed expression and then asked to generate each action and hold it for about five seconds, during which the measurement is done. Patch variances are calculated for the eye closure and grin actions with respect to the relaxed condition, thus enabling it to compare relative movements in facial parts in respective deformations. The quantification of measured variances is carried out by generating a sample population with six normal subjects. The regression coefficient and standard error of the population is computed (fig.8). The measure of quantification is build upon comparing the deviation of population of measured subject with respect to the population of the sample generated with normal subjects.
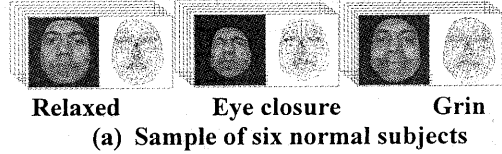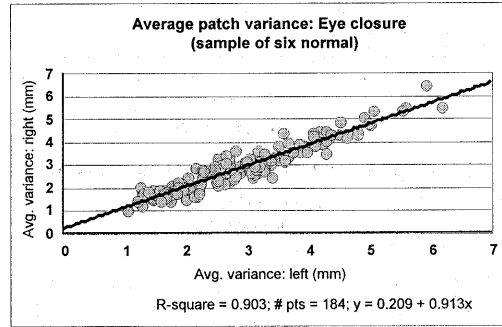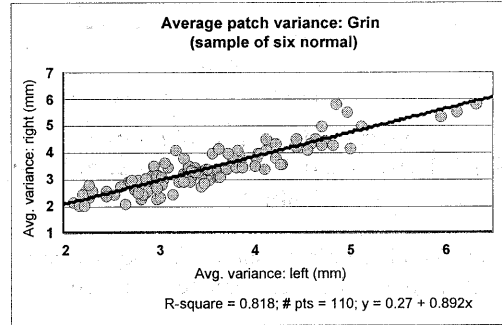

Relaxed   Eye closure   Grin

(a) Sample of six normal subjects

(b) Eye closure with respect to relaxed

(c) Grin with respect to relaxed
Fig.8: Relative patch variances of normal sample.

This distribution with six normal subjects is taken as the control set to determine the deviations in the disordered subjects. Thus, we measure two disordered subjects and plot their distributions in fig.9 and fig.10.
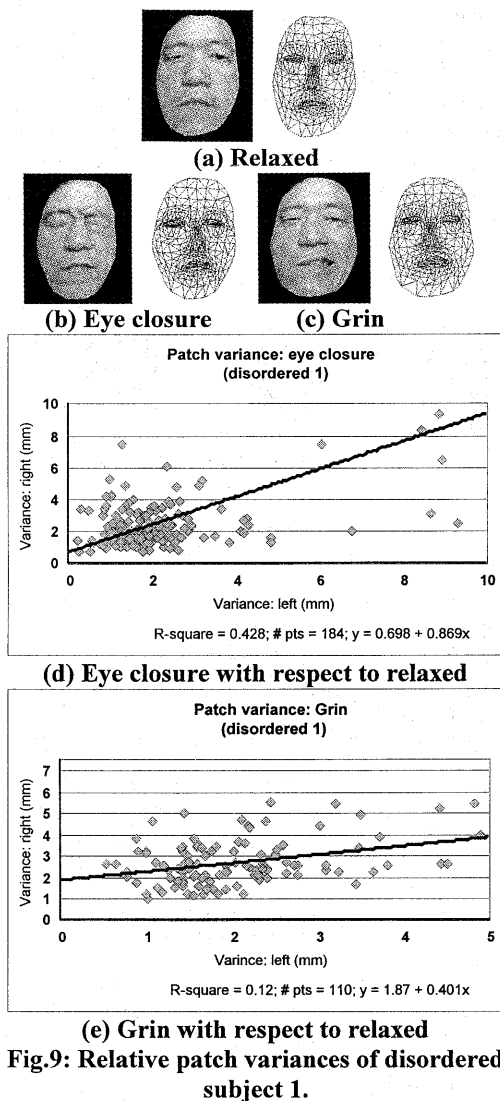
(a) Relaxed

(b) Eye closure   (c) Grin

Patch variance: eye closure
(disordered 1)

R-square = 0.428; # pts = 184; y = 0.698 + 0.869x

(d) Eye closure with respect to relaxed

Patch variance: Grin
(disordered 1)

R-square = 0.12; # pts = 110; y = 1.87 + 0.401x
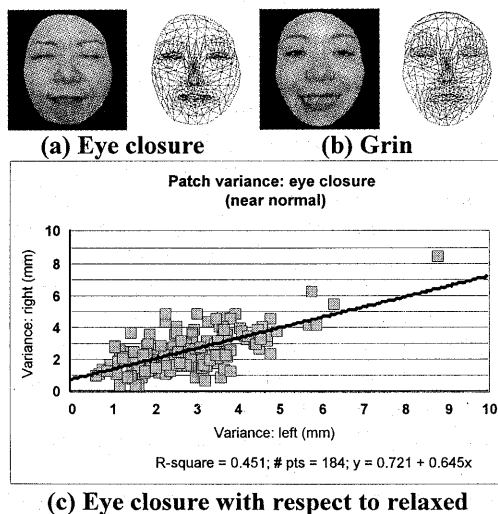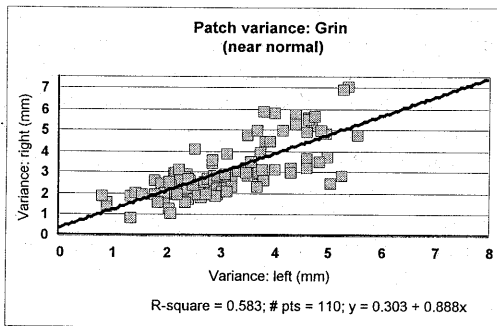
(e) Grin with respect to relaxed
Fig.9: Relative patch variances of disordered
subject 1.

The second disordered subject is also measured
with similar actions and depicted in fig.10.

(a) Relaxed

(b) Eye closure   (c) Grin

Patch variance: eye closure
(disordered 2)

R-square = 0.248; # pts = 184; y = 1.34 + 0.418x

(d) Eye closure with respect to relaxed

Patch variance: Grin
(disordered 2)

R-square = 0.16; # pts = 110; y = 1.51 + 0.382x

(e) Grin with respect to relaxed
Fig.10: Relative patch variances of
disordered subject 2
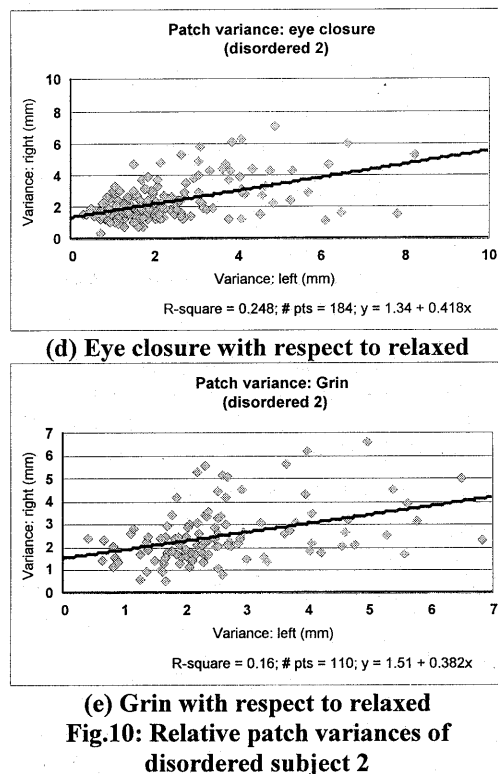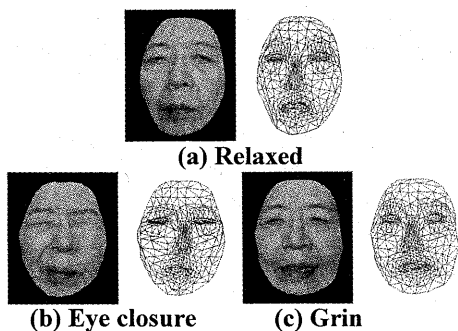
In addition these, we measure a subject who
has almost recovered from nerve paralysis
disorders to evaluate the improvement which is
somewhat difficult to interpret in human
observations. Fig.11 shows the near-normal
subject with similar measurements.

(a) Eye closure   (b) Grin

Patch variance: eye closure
(near normal)

R-square = 0.451; # pts = 184; y = 0.721 + 0.645x

(c) Eye closure with respect to relaxed

(d) Grin with respect to relaxed
Fig.11: Relative patch variance of near normal subject.

The standard error of the populations are computed and depicted in table 1 bellow.
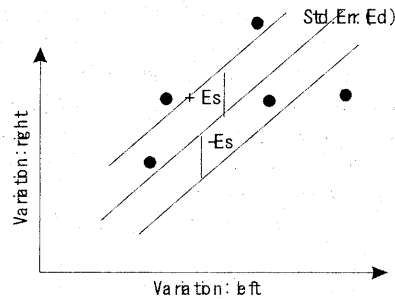
Standard Error		
	Eye-closure	Grin
Sample (6)	0.305	0.357
Near normal	0.856	0.817
Disordered 1	1.850	1.264
Disordered 2	1.026	1.119

Table 1: Standard error of the distributions

The asymmetry of a measured subject is quantified with respect to the similar variation in the normal distribution. The consideration is that, normal distribution fairly indicates the possible variations of expressions about the normal subjects, which can be used as a yardstick to evaluate the asymmetry presented in disordered subjects. Thus we compute the data distribution around the least squares regression line shown in each distribution with the percentage of points lying within multiples of standard error factor. Then taking the standard error of the sample (Es) as a measuring unit, we compute the relative percentage of data lie within the multiples of it in disordered distributions. This phenomenon is depicted in fig.12.
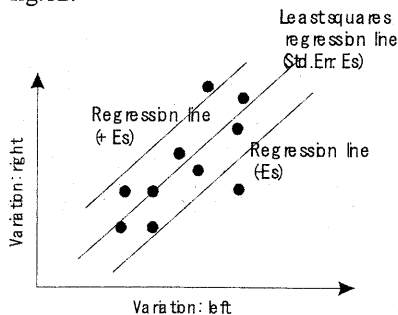
(a) Sample distribution

(b) Disordered distribution
Fig.12: Distribution of points in sample and disordered distributions.

Table 2 shows the percentage of points lied within the multiples of standard error of the normal distribution in all measured distributions for the eye closure action. The same is depicted in Table 3 for the grin action.

Eye closure					
	Within ! Es	Within ! 1.5*Es	Within !2*Es	Within !2.5*Es	Within !3*Es
Sample	69.56	86.41	95.10	97.28	100
Near normal	30.43	42.93	56.52	65.76	75
Dis.sub.1	22.28	32.06	41.84	48.36	58.15
Dis.sub.2	28.80	42.39	53.26	64.13	71.73

Table 2: Point distribution in eye closure

Grin					
	Within ! Es	Within !1.5*Es	Within !2*Es	Within !2.5*Es	Within !3*Es
Sample	71.81	86.36	97.27	98.18	99.09
Near normal	38.18	48.18	59.09	76.36	83.63
Dis.sub.1	26.36	41.81	56.36	65.45	67.27
Dis.subj.2	27.27	39.09	52.72	63.63	71.81

Table 3: Point distribution in grin

5. Discussion

In this work we have demonstrated a technique of estimating asymmetry of a measured facial action with respect to a possible variation of a distribution drawn at normal subjects. We have taken a sample of six apparent-normal subjects and eye closure and grin actions are measured respectively. The average values obtained in this distribution (fig.8(b), 8(c)) are satisfactorily agree with the ideal situation and proved by the small standard error, as depicted in table 1.

In the disordered subjects, the subject 1 shows much asymmetry in both expressions

compared to the subject 2. This is more evident with the high standard error values in its population, compared to the subject 2, which indicates the amount of scatter points presented around its least square regression line. The near normal subject we measured also indicates its improvements compared to the disordered subjects with relatively low standard error (<1.0) factor.

The quantification approach we have proposed, rely on the accuracy of the normal samples measured. It can be seen from table 1, that our normal sample has a low standard error thus, contributing to robust evaluation. The idea behind quantifying disordered distribution with respect to the normal sample is that, it should give a estimate to the evaluators how much the disordered distribution has improved with respect to the similar actions of normal subjects. This is very useful in designing strategies to estimate the amount of facial paralysis in ongoing treatments.

Therefore we quantify the disordered distributions with dividing its regions according to the standard error obtained in the normal samples. According to the tables 2 and 3, we can design a scale to show the amount of asymmetry presented in each measured facial action, with respect to a normal sample.

6. Future work

Our method first adapt a generic face mesh to the measured range data of each subject, and then perform calculations on the generated 3D mesh model. Thus the accuracy of mesh fitting to the measured range data contributes a vital factor for the accuracy of interpretations. Therefore we intend to embed the mesh fitting error factor to the computations by using control subjects with markers attached to the face. The fitting error is then computed by tracking down the marker locations and actual mesh movements in the fitting process.

Further we intend to design a robust scale to represent the asymmetry in each measured action that would be used in clinical practices.

7. Acknowledgement

8. Reference

[1] F.I. Parke, "Computer Generated Animation of Faces", Proceedings, ACM annual conference, pp. 451-457, 1972.

[2] S.M. Platt, N.I.Badler, "Animating Facial Expressions", Computer Graphics, Vol.15, No.5, pp. 245-252, 1981.

[3] M.Isono, K.Murata, H.Tanaka, M.Kawamoto, H.Azuma, "An objective evaluation method for facial mimic motion", Jour. Of Otolaryngology – Head and Neck Surgery, pp.27-31, 1996.

[4] F. Parke. Parameterized model for facial animation. IEEE Computer Graphics and Applications, vol. 2(9), pages 61-68, 1992.

[5] M. Oka, K. Tsutsui, A. Ohba, Y. Kurauchi, T. Tago. Real time manipulation of texture mapped surfaces. Computer Graphics, SIGGRAPH, pages 181-188, 1897.

[6] T. Kurihara, K. Arai. A transformation method for modeling and animation of modeling and animation of human face from photographs. State of art in computer animation, Springer-Verlag, pages 45-57, 1991

[7] M. Nahas, H. Hutric, M. Rioux, J. Domey. Facial image synthesis using skin texture recording. Visual Computer, vol. 6(6), pages 337-343, 1990.

[8] K. Hasegawa, K. Hattori, Y. Sato. A High Speed Face Measurement System. Vision Interface '99, pages 196-202, 1999.