國立交通大學

資訊科學與工程研究所

碩 士 論 文

使用色彩與深度感測器之磁振造影與腦磁波圖自動化對位系統

Automatic Co-registration of EEG-MRI data
using RGB-D Camera

研 究 生：鄭晴懌

指導教授：陳永昇　教授

中 華 民 國 一 百 零 二 年 六 月

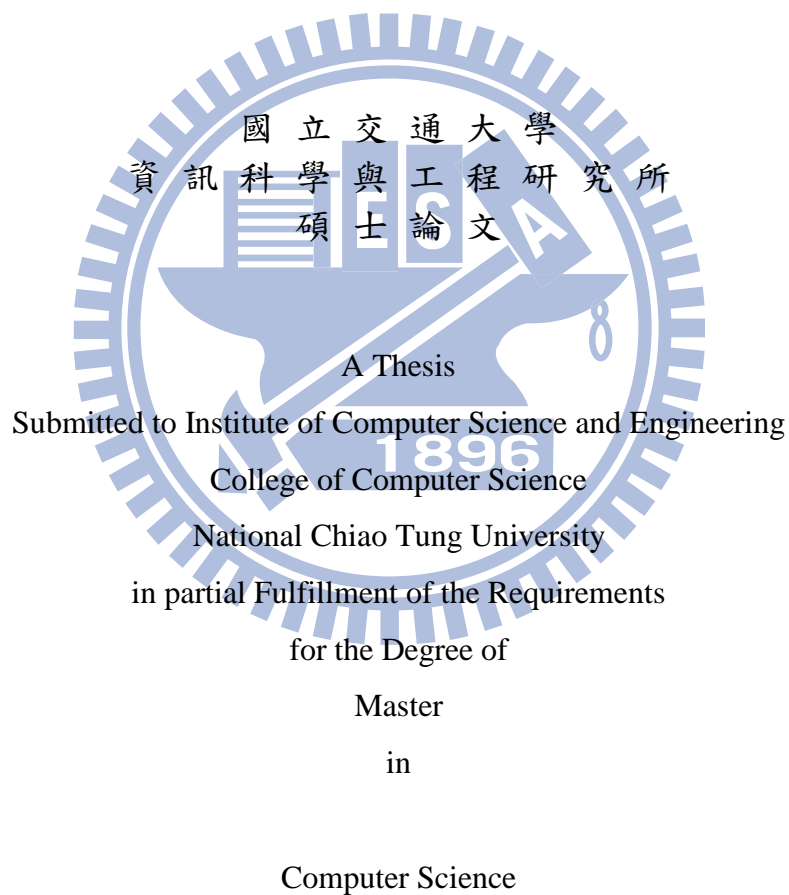使用色彩與深度感測器之磁振造影與腦磁波圖自動化對位系統
Automatic Co-registration of EEG-MRI data Using RGB-D Camera

研 究 生：鄭晴懌　　　　　　　Student：Chiny-Yi Cheng

指導教授：陳永昇　　　　　　　Advisor：Yong-Sheng Chen

國 立 交 通 大 學
資 訊 科 學 與 工 程 研 究 所
碩 士 論 文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2013

Hsinchu, Taiwan, Republic of China

中 華 民 國 一 百 零 二 年 六 月

摘　　要

　　本研究中我們發展一套自動化系統做為磁振造影與腦磁波圖對位。腦磁波圖是一種非侵入式方法提供高時間解析度的腦神經活化訊號但卻缺乏解剖構造資訊，而磁振造影可以提供卓越的空間解析度。因此整合磁振造影與腦磁波圖可以得到兩者的資訊並且用在評估腦部活動源之研究。先前的對位方法需要額外的標記點及易受到操作者因素影響等缺點，為解決這些問題我們使用 Kinect for Windows 作為介面，將磁振造影座標系統和數位筆(digitizer)座標系統對位在一起。

　　本研究方法主要分成兩個部分，第一部分為使用臉部資訊將 Kinect 與磁振造影的座標系統對位在一起，第二部分為利用臉上事先貼好的標記點在 Kinect 座標系統和數位筆座標系統中的相對應座標找彼此的轉換關係。經由我們的對位系統，數位筆所點的電極位置會和磁振造影在同一個座標系統。由於記錄電擊帽上的電極位置並不是貼緊在頭皮上面，我們利用厚度補償來調整電極的位置。最後使用三種評估方法來驗證對位的結果並比較先前所作的方法。

　　本實驗中有十位受測者參與且每位做了兩次實驗。所使用的三種評估方法其中平均感測器對位誤差為 1.67 釐米且平均十字對位誤差為 1.83 釐米。比較其他論文的結果顯示我們的系統有足夠的準確性、重複使用性、快速等特性用於磁振造影與腦磁波圖對位。

Abstract

In this study, we developed an automatic system to co-register EEG and MRI data that used for the source estimation of electric brain activity. Integration of EEG and MRI brain data could obtain high temporal and spatial resolution that can localize sources of rapid changes of cortical activation. In the conventional method, there are still some drawbacks such as the need of additional markers and the manual error. Thus, by using Kinect for Windows which extracts color and depth information at the same time as interface, we aligned MRI coordinate system with digitizer coordinate system to resolve problems.

In the proposed method, a complete EEG-MRI co-registration would be achieved in two steps: (i) Aligned Kinect coordinate system with MRI coordinate system using facial information. (ii) The transformation is based on matching of corresponding markers pasted on the face during acquisition of both in the Kinect and digitizer coordinate systems. By our system, the location of electrodes which were recorded by digitization device would be found in the same coordinate system. Due to the thickness of electrodes on the EEG cap, we adjusted them with thickness compensation. Finally, using three error estimations to evaluate the results of co-registration and compared with different methods made before.

The experiment was performed with ten subjects, and each one done twice. The mean residual error of sensors was 1.67 mm, and the mean residual error of cross was 1.83 mm. These results compared with other methods show that our system is sufficiently accurate, repeatable, efficient and labor-intensive to be used to assist neuroscience studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

The goal of brain source images is to investigate the brain activity which can be estimated by measuring electroencephalography (EEG) signals and solving inverse problem. In which EEG provides excellent temporal resolution noninvasively but lacks for anatomical information. However, Magnetic Resonance Imaging (MRI) is also noninvasive method but provides anatomical structure of brain. Therefore, co-registration of EEG and MRI data that can combine the functional and structural information is profitable to localization and observation of the brain activity over the subject's cortex.

Generally, the co-registration of EEG and MRI data includes two steps. First, we can know the positions of EEG sensors in the three-dimensional space by any digitization devices. Second, by co-registration method that transforms the positions of sensors from digitizer coordinate system to MRI coordinate system, we can localize the brain activity in MRI coordinate system. Figure 1.1 illustrates the co-registration of EEG and MRI data. Accurate anatomical localization of functional brain depends on the precise co-registration, otherwise it will influence on next steps including forward modeling or inverse problem.

MRI coordinate system   **digitizer coordinate system**
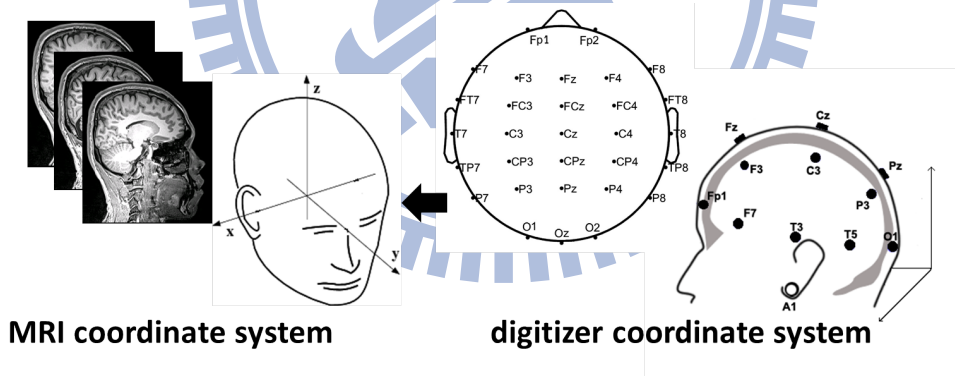
Figure 1.1: The co-registration of EEG and MRI data. By co-registration method that transforms the positions of sensors from digitizer coordinate system to MRI coordinate system.

In the conventional method, the co-registration method required to manually determine at least three landmark positions on the head of subject, and which had to find the same positions both in MRI and digitizer coordinate systems. [15] This method would produce

differences from variable operators or times. Another method is using external fiducial markers visible in MRI without the judgement of operators to find the same positions in two different coordinate systems. [14] But this method was discomfort for participants due to wearing fiducial markers. Besides, the relatively larger markers produce the larger error from transformation. Therefore, some techniques were developed to improve these drawbacks, such as using visible electrodes in MRI [10, 17] and distance-based method. [9, 11, 12, 16] Using visible electrodes in MRI no need to the procedure of alignment, so this method do not have error of transformation between two different coordinate systems. However, due to having to a new MRI scanning every time when experimenting, using MRI visible markers was lack of flexibility and economic benefits, and also not suitable for retrospective studies. In addition, it remains to be proven whether it works with the environment of MRI scans. There was another way that utilizes the digitization device to record over five hundreds points on face and minimize the distance between those facial points and the scalp of head from MR images to find the best transformation relation, yet this cost too much time and labor on recording facial points and the degree of distribution which would affect the final result. Moreover, the different operators also affect the result of registration.

In order to tackle the problems mentioned above, our study used Kinect for windows that could extract the depth and color information at the same time as interface to co-register between MRI and digitizer coordinate systems. In this way, we could find these points recorded by the digitization device precisely through coordinate transformation in MR images. Our method did not need additional markers and more effort of operator, and the result of alignment would not be affected by uneven distribution of these points. Certainly, the time spent in recording the facial points was saved. Furthermore, medical resources would be saved because every subject only had to accept MR scans once.

1.2   Related work

For several years, there are many researches about co-registration of EEG and MRI data. In this section, we would introduce these related works about distance-based method.

Estimation of the accuracy of a surface matching technique for registration of EEG and MRI data

The system aligned head surface which recorded by 3D ultrasound localizing device with segmentation of MRI data. By calculating the centroid of digitized head surface and MRI-derived head surface matched together to give the good initial. The surface matching technique aims to minimizing the cost function that converges to a minimum. And the cost function can see equation 1.1. The accuracy was estimated by calculating the residual error of the surface and intra-subject and inter-subject variability. Figure 1.2 shows the result of registration of EEG and MRI data. [9]

$$\frac{1}{n}\sum_{i=1}^{n} d\left(\mathbf{T}\left(P_i\right)\right) \tag{1.1}$$

where d is the relative distance, P_i (i = 1,2,...,n) the 3D-scanned head-surface points, and $\mathbf{T}$ the rigid transformation that maps the 3D-scanned head-surface points into the coordinate system of the 3D image.

Automatic alignment of EEG/MEG and MRI data sets

This system uses distance-based alignment (See figure 1.3) that minimizes the Euclidean distance transform, and using 3D geometrical moments for the initial alignment. This system was tested on inhomogeneous digitized points.(See figure 1.4) And the results showed the most uniform distribution of points have better results. The average distance between points and the MRI head surface is calculated as error estimation. [12]
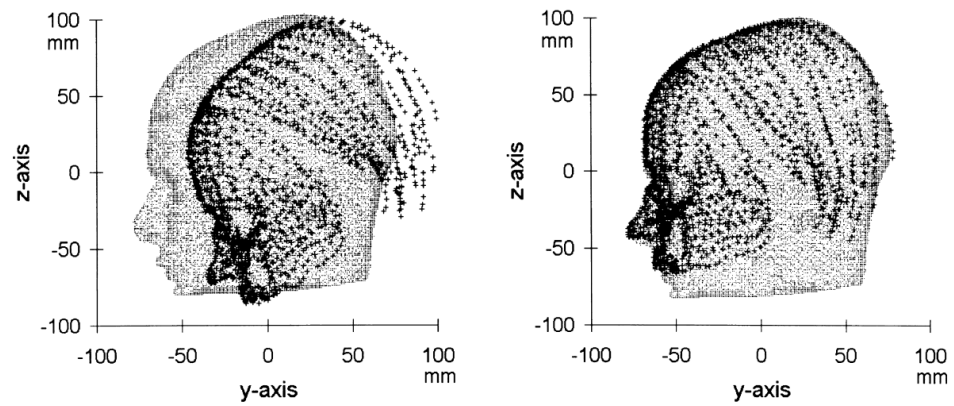
Figure 1.2: The result of registration of EEG and MRI data. Digitised points of the head surface superimposed on the points of MRI-derived head surface before (left) and after (right) surface matching. (Figure source: [9])
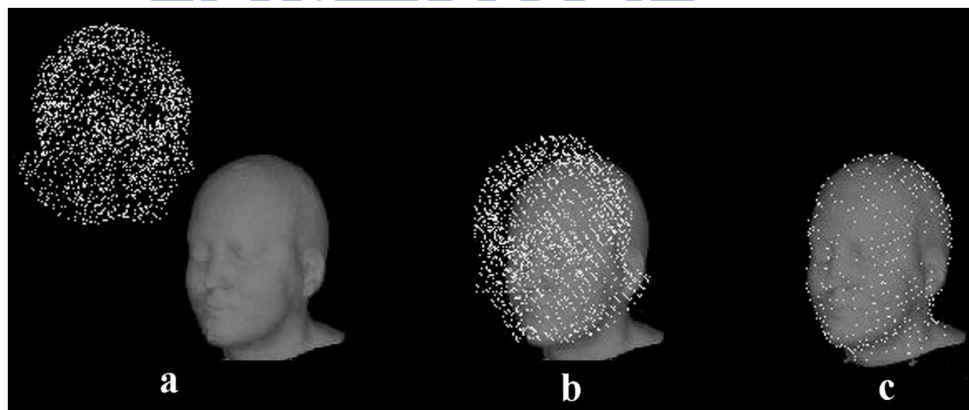
Figure 1.3: The concept of distance-based alignment. (Figure source: [12])
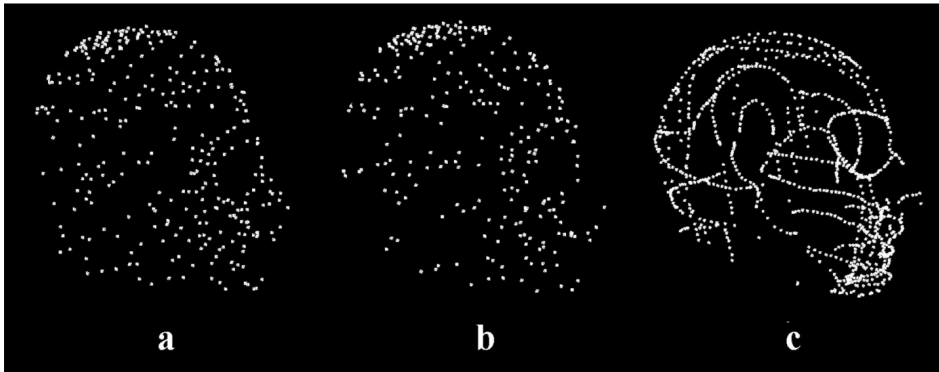
Figure 1.4: The inhomogeneous digitized points. Subject a: the most uniform distribution of points. (Figure source: [12])

Validation of a method for coregistering scalp recording locations with 3D structural MR images

The segmentation of MRI scalp points was chosen from the out to inside until the threshold was reached and the user could adjust the threshold value to extract the MRI scalp points. The sensor locations and facial points were obtained with a Polhemus. The MR-derived points are registered to the digitized points recorded in the coordinate system. And the algorithm uses a MarquardtLevenberg optimization routine that minimizes the sum of squares of the distances between the two data sets. After alignment, this study used two non-rigid procedures: (i) scaling that using independent linear scaling factors to reduce registration errors and (ii) scalp forcing that replacing the digitized points with the closest segmented MR scalp points. Figure 1.5 denotes the misalignment and the result of co-registration. [16]

EEG-MRI Co-registration and sensor labeling using a 3D laser scanner

This system acquires the positions of EEG sensors and the facial points by using 3D handheld laser scanner (see figure 1.6), and the scalp points was obtained from the MR images. A pre-alignment was given by using three ducial landmarks for co-registration. As to co-registration, this system used iterative closest point algorithm that minimize the distance between the two sets of points. This study uses two error estimations which are

<div align="center">(a) (b)</div>
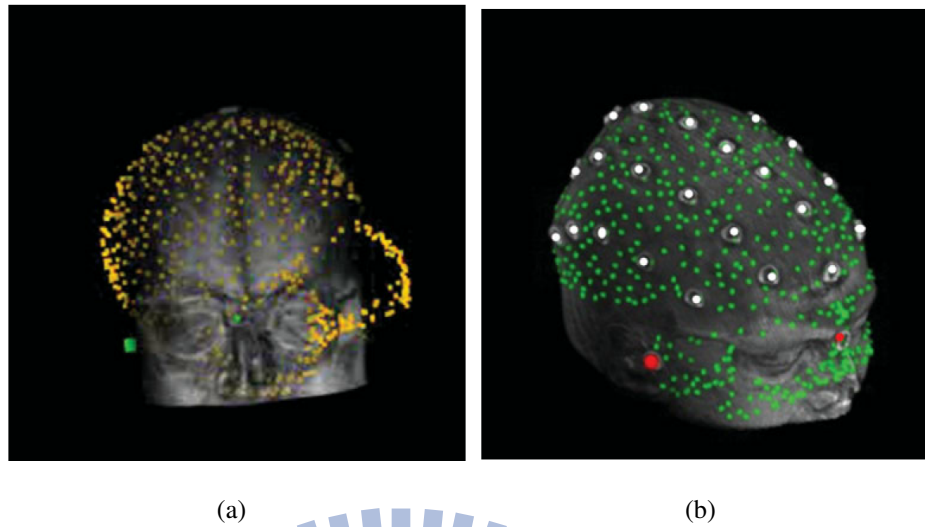
Figure 1.5: The misalignment and the result of co-registration. (Figure source: [16])

residual error of the face and residual error of the sensor to evaluate the accuracy of co-registration results. Figure 1.7 shows the result of EEG-MRI co-registration and sensor labeling. [11]
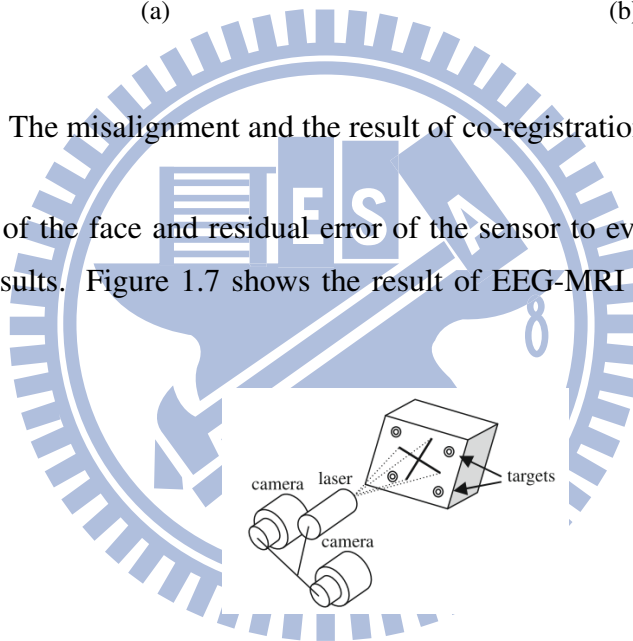
Figure 1.6: The 3D handheld laser scanner. (Figure source: [11])
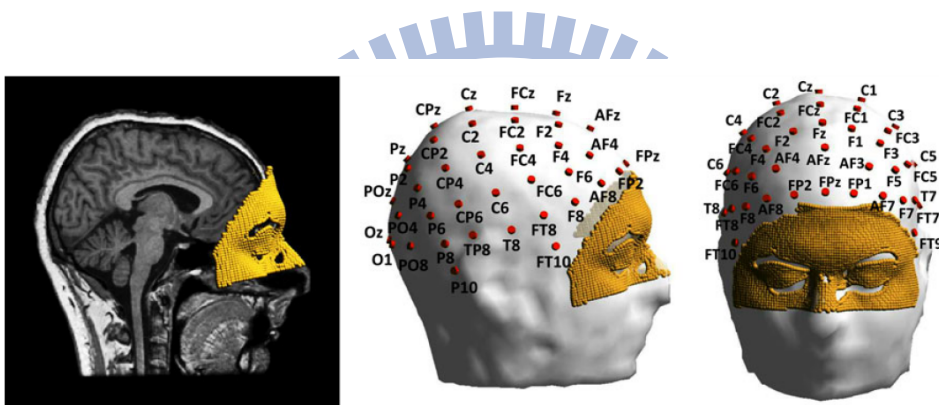
Figure 1.7: The result of EEG-MRI co-registration and sensor labeling. The MRI and headmodel co-registered with the scanned face and the positions of sensor. [11])

1.3 Kinect for windows

Kinect is a widely available, inexpensive device that provides color, depth and audio information. Depth data is obtained by using light coding, which mainly utilizes laser speckle to code the objects in the space. Kinect sensor projects infrared light uniformly distributed in the environment through IR emitter. When the rough object is irradiated with infrared light, it will form random reection spots, called speckle. The encoding light is read by the CMOS sensor and decoded by the chip to create a depth map. Figure 1.8 shows the basic operations of Kinect sensor.
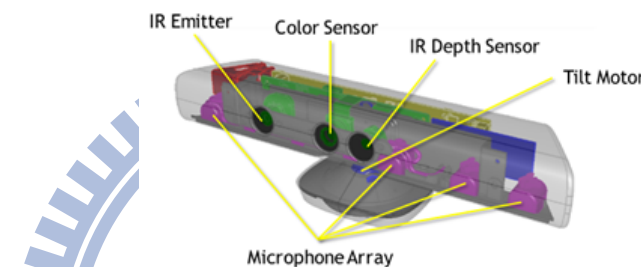
Figure 1.8: Basic operations of Kinect for windows. (Figure source: [2])

Microsoft released latest Kinect on February 1st, 2012, called "Kinect for windows", which was designed to connect to the computer that makes inventors able to programs about Kinect under Windows. In addition to those functions of previous Kinect Xbox360 has, there are more update be added. The most noticeable function is that it provides "Near mode", which enables to view objects near to 40 centimeters in front of the device through depth camera accurately. Furthermore, these changes make the sensor less sensitive to far distance objects. When the sensor is in near mode, it is possible to view objects two meters away precisely, but out to three meters the depth value is unstable. Seeing figure 1.9 that show the visible depth values.

There is specification of Kinect for windows, seeing table 1.1.
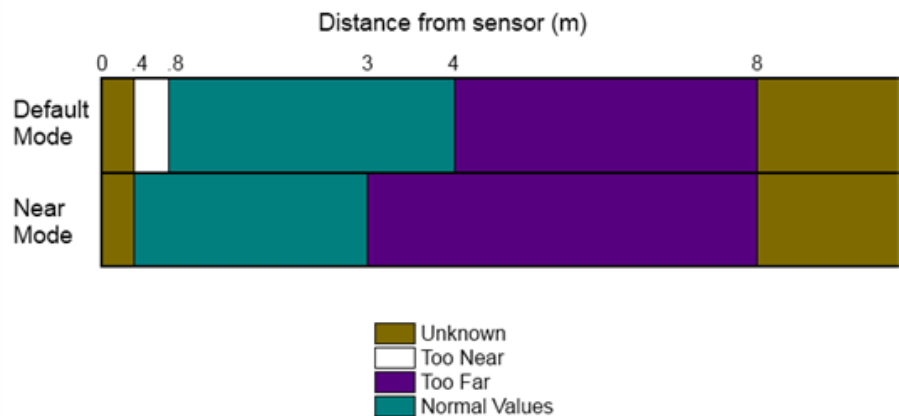
Figure 1.9: The depth values that Kinect for windows can view. (Figure source: [2])

Table 1.1: The specification of Kinect for windows.

Sensor	Color and depth-sensing lenses
	Voice microphone array
	Tilt motor for sensor adjustment
Field of View	Horizontal eld of view: 57 degrees
	Vertical eld of view: 43 degrees
	Physical tilt range: +-27 degrees
	Depth sensor range: 0.8m - 4m
	Near mode Depth range: 0.4m 3m
Data Streams	320x240 16-bit depth @ 30 frames/sec
	640x480 32-bit color@ 30 frames/sec
	16-bit audio @ 16 kHz
Skeletal Tracking System	Tracks up to 6 people, including 2 active players
	Tracks 20 joints per active player
	Ability to map active players to Xbox LIVE Avatars
Audio System	Echo cancellation system enhances voice input
	Speech recognition in multiple languages

1.4   Thesis overview

Figure 1.10 shows the concept of our system. Our system used Kinect for windows that could extract the point could which owns depth and color information at the same time as interface to co-register between MRI and digitizer coordinate systems. At first, we extracted points on facial surface from MR images and aligned with facial points which extracted by Kinect for windows. Second, using the coordinate correspondences of markers to find the relation between digitizer and Kinect coordinate systems. Then we could localize the EEG sensors in the MRI coordinate system.

MRI coordinate system   **Kinect coordinate system**   **digitizer coordinate system**

Figure 1.10: The concept of our system.

1.5   Thesis organization

The remaining part of this thesis is organized as follows. Chapter 2 presents the methods used in our system. The experiment setup and the results are given in Chapter 3. Some discussions related to our methods are in Chapter 4 and Chapter 5 is conclusions.

Chapter 2

Methods

2.1 Overview

At first, the subject would be pasted on four markers and wear head-sized EEG cap. Apply Kinect for windows that is able to extract depth and color information closely as interface to extract point cloud from the view of Kinect for windows. After that we used depth threshold to segment the head points from the background scene and do pre-processing in order to extract the facial points without any noise or outlier. Next, we used scalp segmentation to extract points on facial surface from MR slices and then align with Kinect-derived facial points.

On the other hand, apply the digitization device to record the centroid of markers on face and electrodes, then using color information to find the region of markers and calculate the locations of centroids of markers automatically. In this way we found the transformation relation between digitizer and Kinect coordinate systems by using the two coordinate correspondences. However, due to the thickness of electrode on EEG cap, we adjusted them with the thickness compensation. Finally, we could find these points recorded by the digitization device precisely in MRI slices through coordinate transformation. The main structure of our system shows in Figure 2.1.

This chapter would be divided into the following: (i) segmentation of scalp points from MRI; (ii) extraction of RGB-D data; (iii) depth threshold; (iv) pre-processing; (v) co-registration; (vi) estimation of centroid of markers; (vii) estimation of transformation relation; (viii) error estimation.

2.2 Segmentation of scalp points from MRI

We used the gray value threshold to find the border between scalp and outer surrounding air from left-right, right- left, anterior-posterior and up-down directions, and the union of the points from all scanning directions as the scalp points (S). According to [16], the gray value threshold was defined as a fraction of the largest gray value at first (threshold = 0.04). In this procedure, the operator was allowed to adjust the threshold manually, in order to find the correct scalp points. The operator could obverse the segmentation of scalp points which is smooth enough or not (expose the inner structure of brain) to decide whether to
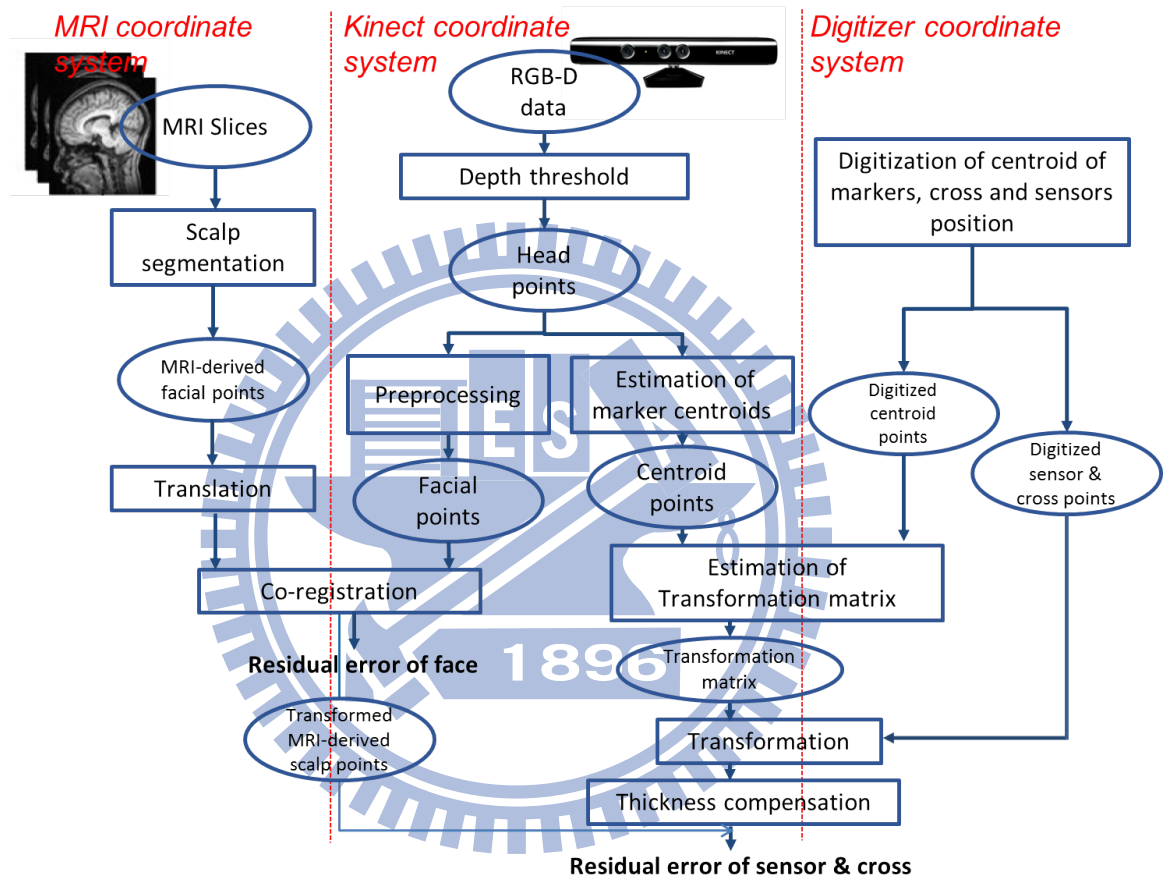
Figure 2.1: Flow chart of automatic co-registration of EEG-MRI data using RGB-D camera.
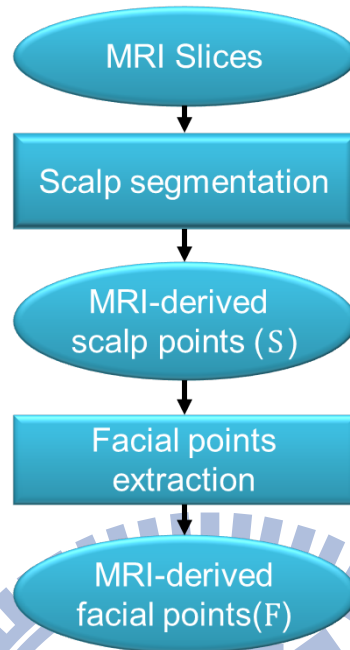
Figure 2.2: Flow chart of segmentation of scalp points from MRI.

adjust the threshold. After segmentation done, we could get about 80,000 140,000 all head scalp points, and we extract about front one-third (rough 40,000 -60,000 points) to obtain the points of facial region (F) to align with facial points extracted by Kinect for windows in order to reduce the time cost for the registration and increase the accuracy of the registration result. The flow chart can see Figure 2.2.

2.3 Extraction of RGB-D data

We extracted the data of Kinect for windows by OpenNI [4] and utilized the transformation formula of OpenNI to get the point and its coordinate value of x, y, z. In addition, OpenNI provides Calibration function to correct the visual angles between two cameras and align the object to the same place. We required the light of space be brightened and balance to extract RGB-D data and used color information to estimate the accuracy of marker centroids.

In order to acquire the more accuracy facial points extracted by Kinect for windows,
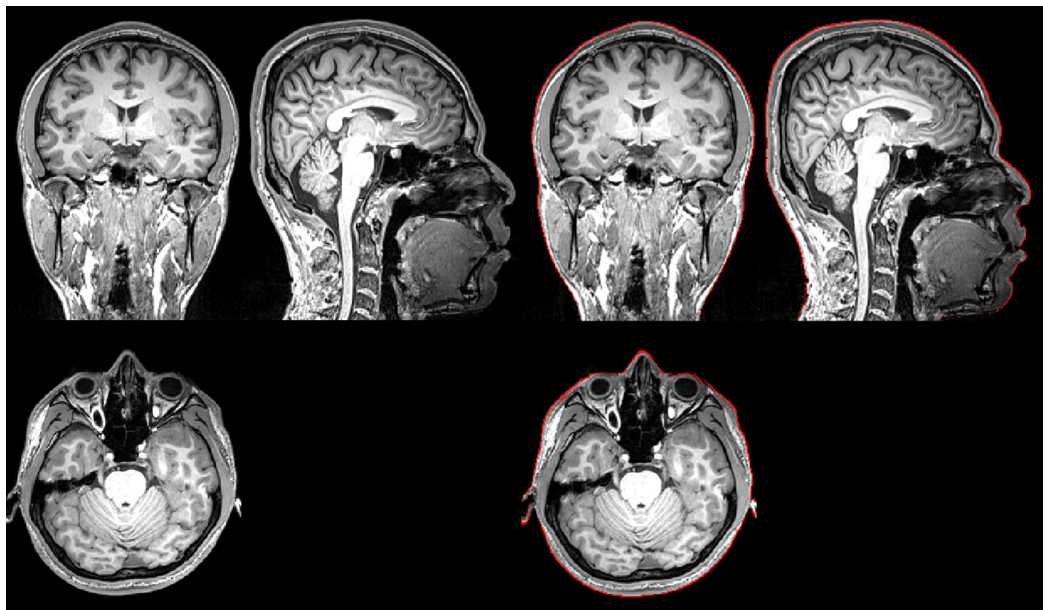
(a)
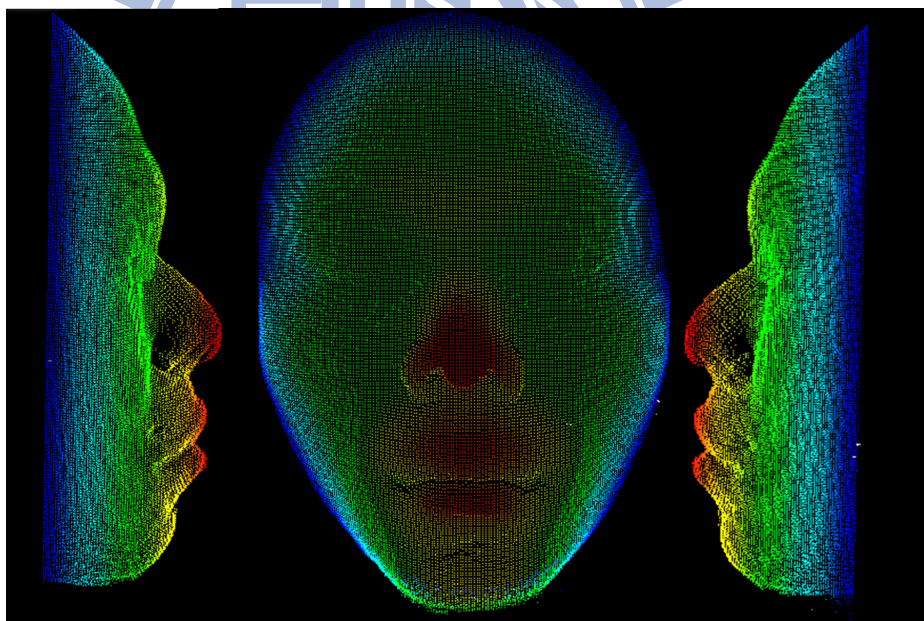
(b)

Figure 2.3: The result of segmentation of Scalp points from MRI. (a) The left shows the MR image, and right shows the segmentation of scalp points that superimposed on the MR image. (b) The facial points are viewed from three angles.

the subject have to near the Kinect sensor as close as possible. However, when the subject close to about 0.46 meter, some parts depth information such as tip of nose would be unstable. Sometimes Kinect could receive depth information at there, but sometimes could not. We had to reduce these missing data before further processing or have the subject move backward for a short distance. According to [8], we averaged thirty depth images in real-time streaming to reduce noise and missing data.

2.4 Depth threshold

At first, we could use depth threshold to separate the head point cloud(H) from data set due to the subject sat closely in front of the Kinect for windows. As to threshold value, we found the range between the smallest depth value in the point cloud and back extension about 0.15 meters as head points. The equation can see below:

$$H = \{x | x_z < \min_z + 0.15\,\text{m}, x \in \text{points extracted by Kinect.}\} \tag{2.1}$$

2.5 Pre-processing

We wanted facial points to align with the MRI-derived facial points without any noise or outlier to affect the result of registration. But after depth threshold, there still have points of EEG cap and markers region in the Kinect-derived head points. So we used color information to extract the points of facial region, and used some methods to remove outliers and noises. As below we partitioned into (i) skin color filter and (ii) outlier removal to explain how to get clean Kinect-derived facial points. Figure 2.5 shows the result of preprocessing.

2.5.1 Skin color filter

We could take advantage of color information to remove the part of EEG cap and makers from point cloud. We transformed the color information of point cloud from RGB to YCbCr color space.(See equation 2.1) According to Chai and Ngan [7] mentioned the

skin color region, we obtained the facial region by extracting those skin color range in the YCbCr space, and the skin range could see equation 2.2.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.168 & -0.331 & 0.5 \\ 0.5 & -0.418 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} \tag{2.2}$$

$$\text{Skin} = \begin{cases} 1, \text{ if} \begin{cases} 77 < Cb < 127 \\ 133 < Cr < 173 \end{cases} \\ \quad 0, \text{ otherwise} \end{cases} \tag{2.3}$$

2.5.2 Outlier removal

Kinect for windows typically generate point cloud of varying point densities. Additionally, measurement errors lead to sparse outliers which corrupt registration result even more. Some of these irregularities could be solved by performing a statistical analysis on the neighbour of each point, and removing those points which do not have a certain criteria. Some of points are dispersed like outliers or noise so we remove the points which do not have enough number of neighbuor points within a certain range.

Statistical outlier removal

The concept of statistical outlier removal is based on the computation of the distribution of distance between each point and its neighbours in the data set. For each point x, we calculated the mean distance d between it to its two hundred neighbour points.

$$d(x) = \underset{y \in kNN}{\text{avg}} (\|x - y\|_2) \tag{2.4}$$

Following we calculated the global mean $\bar{\text{d}}$ and standard deviation SD from the mean distance of all points in the data set. By assuming that the data set was Gaussian distribution, the point in the data set would be considered as outlier while its mean distance was larger than the global mean plus one standard deviation and would be removed. [5]

$$\left| d(x) - \bar{d} \right| > 1\text{SD} \tag{2.5}$$

Radius outlier removal

The concept of radius outlier removal is that removing the point which does not have enough number of neighbour points within a certain range. We set a number of neighbour points which each point must have within a certain radius in the point set. Figure 2.4 shows what the radius outlier removal filter does. For example if one neighbor point is specified, the star point will be removed. If two neighbor points are specified then both the star and triangle points will be removed. [5]
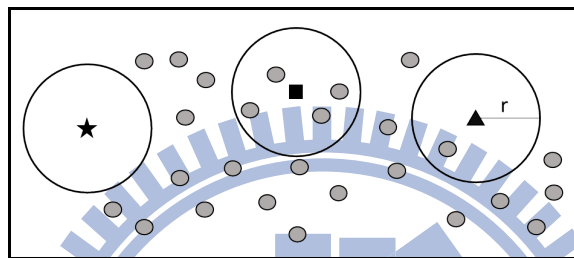
Figure 2.4: Schematic diagram of radius outlier removal. If threshold T is one, then star point will be removed. If threshold T is two, then star and triangle points will be removed.

Erosion

Owing to the definition of skin color range is overly broad, the region connected with the edge of face such as the brim of EEG cap would be extracted whereas it was not the region of skin. Therefore, operators may make a judgement about the point cloud they extracted to eliminate the edge of face by erosion.

2.5.3 Down-sampling

In order to speed up the registration between Kinect and MRI coordinate systems using facial points, we reduced the Kinect-derived facial points uniformly to do registration. The method of down-sampling is giving threshold T and choosing the pixel at intervals of T both in the horizontal and vertical directions. Operator can adjust the threshold to control number of facial points to do co-registration. Figure 2.6 represents the different level of down sampling facial points.
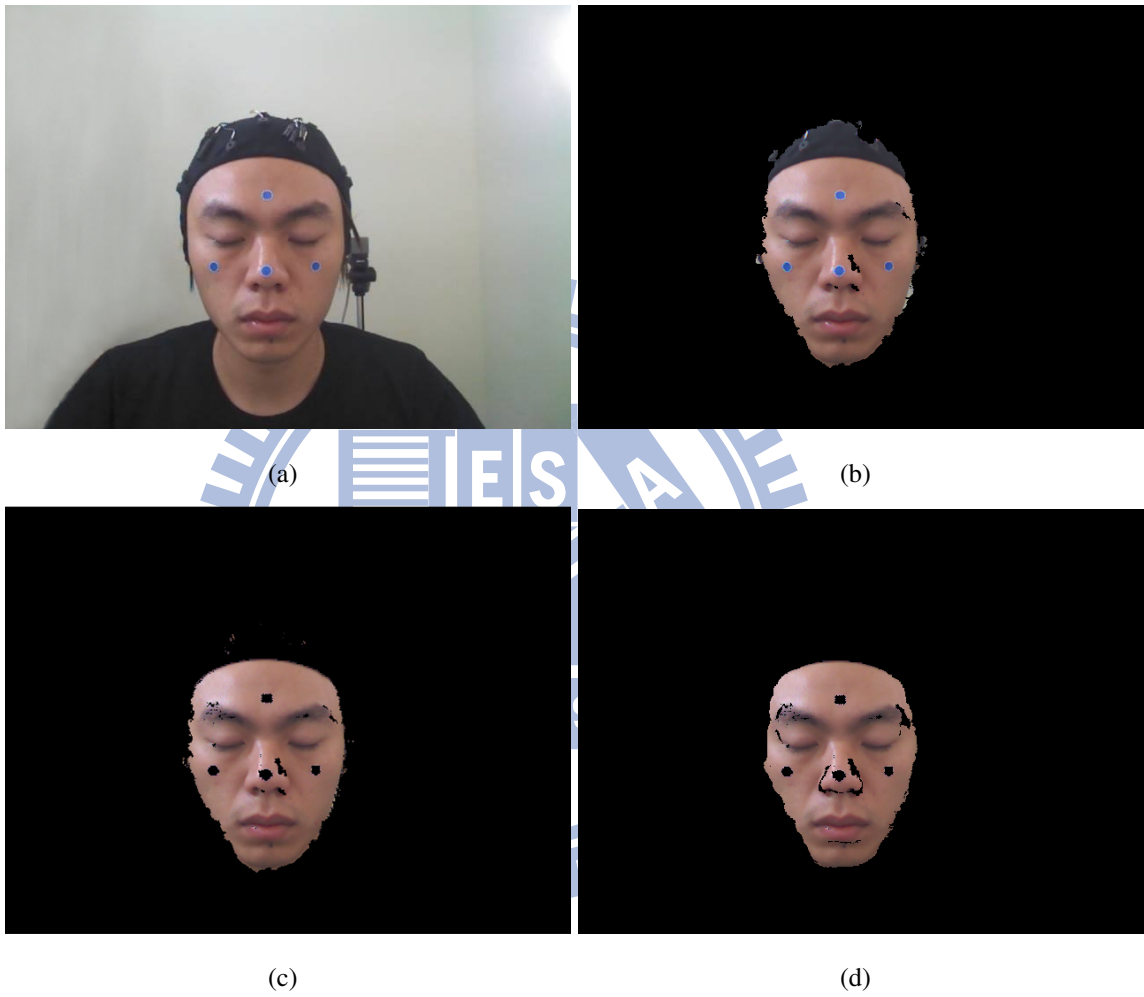
Figure 2.5: The results of preprocessing. (a) A original image. (b) After depth threshold. (c) After skin color filter . (d) After outlier removal.

2.6   Co-registration

The conventional ICP algorithm [6] is aim to iterative a group of data set to another until the Euclidean distance of two data sets was minimal. The equation can see below:

$$\left(\hat{\mathbf{R}}, \hat{\mathbf{t}}\right) = \arg \min_{(\mathbf{R}, \mathbf{t})} \sum_{j=1}^{M} \sum_{i=1}^{N} \left(\|\mathbf{x}_i - \mathbf{R}\mathbf{p}_j - \mathbf{t}\|^2\right) \tag{2.6}$$

where $\mathbf{t}$ is translation vector, $\mathbf{R}$ is Rotation matrix, $\mathbf{x}_i$ and $\mathbf{p}_j$ are points in two data sets respectively for $i = 1$ to N and $j = 1$ to M.

However, it needed a good initial value and tend to result in local minimal to get poor registration effect. For our subsequent registration, the centroid of mass of the Kinect-derived facial points and the MRI-derived facial points are matched to give an appropriate initial value. We use modified coherent point drift (CPD) algorithm [13] to do registration.

2.6.1   Modified coherent point drift algorithm

Unlike ICP, Coherent Point Drift Algorithm uses soft assignment of correspondences that establishes correspondences between all combinations of points. Coherent point drift algorithm consider the alignment of two data sets as a probabilistic method, and considering the alignment of two data sets as a probability density estimation problem, where we fit Kinect-derived facial points as the Gaussian Mixture model (GMM) centroids and MRI-derived facial points as the data points. At the optimum, two data sets become aligned
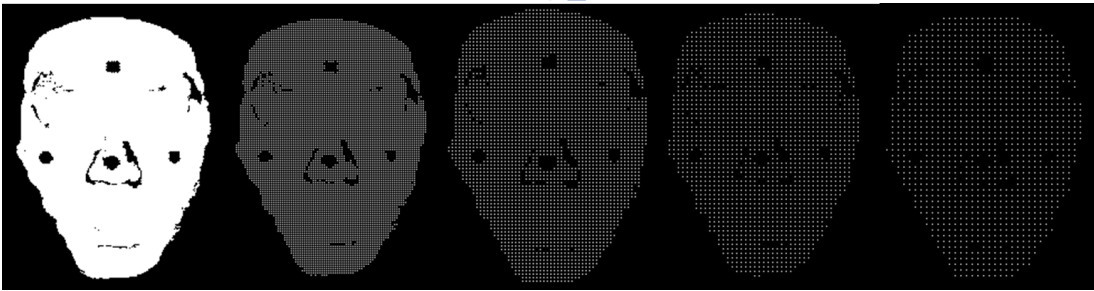
Figure 2.6: The different level of down sampling facial points. The leftmost is original facial points, and the number of points are 29,770. The left two are 7,443. In the middle are 3,296. The right two are 1,855, and the rightmost are 1,185.

and the correspondence is obtained using the maximum of the GMM posterior probability for a given data set. And Andriy et al. [13] derive a closed form solution of the maximization step of the EM algorithm. The algorithm can see below: Since Kinect-derived face points contain slightly noise and outliers, we had to reduce the poor effect of registration and speed up computing. We added considering the average minimal distance of Kinect-derived faces points that were larger than three standard deviations as outliner in the iterative process. The outliers were not included in the registration process to increase the speed of convergence. After registration, the MRI-derived scalp points were aligned with the Kinect-derived face points in the same coordinate system and the transformation relation was known.

2.7 Estimation of centroid of markers

At first, we converted point cloud which has color information from RGB color space to HSL color space (See figure 2.9). HSL is an acronym for hue-saturation-lightness, and the value of hue would not be changed by the different lightness we photo every time, so we adopted this color space to extract blue region. By testing the range of hue, we could extract the markers easily from point cloud. We extract markers region by subtracting the histogram with a predefined face-only histogram. (Seeing figure 2.11(b) the red block.) The range of hue values would include the white color, and next made out the blue region and remove white color by analysis of histograms from red and green domains. We found the clear boundaries to distinguish between blue and white in red domain and green domain.
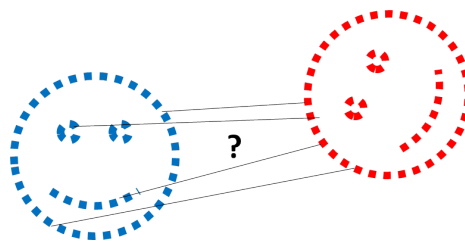
Figure 2.7: Schematic diagram of registration. Given two point sets, assign the transformation relation that maps one point set to the other.

Rigid point set registration algorithm:

- Initialization: $\mathbf{R} = \mathbf{I}, \mathbf{t} = 0, s = 1, 0 \leq w \leq 1$

  $\sigma^2 = \frac{1}{DNM} \sum_{n=1}^{N} \sum_{m=1}^{M} \|\mathbf{x}_n - \mathbf{y}_m\|^2$

- EM optimization, repeat until convergence:

  - E-step: Compute $\mathbf{P}$,

    $p_{mn} = \dfrac{\exp^{-\frac{1}{2\sigma^2}\|\mathbf{x}_n - (s\mathbf{R}\mathbf{y}_m + \mathbf{t})\|^2}}{\sum_{k=1}^{M} \exp^{-\frac{1}{2\sigma^2}\|\mathbf{x}_n - (s\mathbf{R}\mathbf{y}_k + \mathbf{t})\|^2} + (2\pi\sigma^2)^{D/2}\frac{w}{1-w}\frac{M}{N}}$

  - M-step: Solve for $\mathbf{R}, s, \mathbf{t}, \sigma^2$:

    · $N_{\mathbf{P}} = \mathbf{1}^T\mathbf{P}\mathbf{1}, \mu_{\mathbf{x}} = \frac{1}{N_{\mathbf{P}}}\mathbf{X}^T\mathbf{P}^T\mathbf{1}, \mu_{\mathbf{y}} = \frac{1}{N_{\mathbf{P}}}\mathbf{Y}^T\mathbf{P}\mathbf{1}$,

    · $\hat{\mathbf{X}} = \mathbf{X} - \mathbf{1}\mu_{\mathbf{x}}^T, \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{1}\mu_{\mathbf{y}}^T$,

    · $\mathbf{A} = \hat{\mathbf{X}}^T\mathbf{P}^T\hat{\mathbf{Y}}$, compute SVD of $\mathbf{A} = \mathbf{U}SS\mathbf{V}^T$,

    · $\mathbf{R} = \mathbf{U}\mathbf{C}\mathbf{V}^T$, where $\mathbf{C} = \mathrm{d}(1, .., 1, \det(\mathbf{U}\mathbf{V}^T))$,

    · $s = \dfrac{\mathrm{tr}(\mathbf{A}^T\mathbf{R})}{\mathrm{tr}(\hat{\mathbf{Y}}^T \, \mathrm{d}(\mathbf{P}\mathbf{1})\hat{\mathbf{Y}})}$,

    · $\mathbf{t} = \mu_{\mathbf{x}} - s\mathbf{R}\mu_{\mathbf{y}}$,

    · $\sigma^2 = \frac{1}{N_{\mathbf{P}}D}(\mathrm{tr}(\hat{\mathbf{X}}^T \, \mathrm{d}(\mathbf{P}^T\mathbf{1})\hat{\mathbf{X}}) - s\,\mathrm{tr}(\mathbf{A}^T\mathbf{R}))$.

- The aligned point set is $\mathcal{T}(\mathbf{Y}) = s\mathbf{Y}\mathbf{R}^T + \mathbf{1}\mathbf{t}^T$,
- The probability of correspondence is given by $\mathbf{P}$.

Figure 2.8: Coherent point drift algorithm. [13]

Figure 2.11(d)(e) shows the histogram of red domain and green domain.
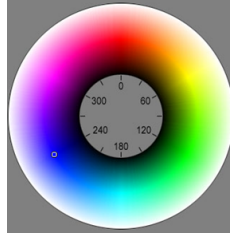
Figure 2.9: HSL color space. (Figure soure: [1])

After using color information to extract the points of blue region, we used connected component to distinguish which points belong which regions from the point cloud, and calculate the respective centroids of regions by averaging their coordinate values. Finally, we could get the coordinate value of the markers' centroid. Figure 2.11 shows the result.

2.8   Estimation of transformation relation

In the experiment procedure, we have digitized the centroid of markers that stuck on the subjects face, and we could get the coordinate value of the points after finding the centroid by our procedure in the Kinect coordinate system. We took advantage of their correspondence between the coordinate values of each other point to find the transformation matrix. Our approach used least square method that minimizing the sum of squared differences between the corresponding positions. The equation can see below:

$$\hat{\mathbf{T}} = \arg \min_{\mathbf{T}} \left(\sum_{i=1}^{N} \left(\|\mathbf{x}_i - \mathbf{T}\mathbf{p}_i\|^2 \right) \right) \tag{2.7}$$

where $\mathbf{x}_i \in$ centroid points, and $\mathbf{p}_i \in$ digitized centroid points for i =1 to N.

First, setting those coordinates values of corresponding points in respective matrices. Second, calculating the pseudo-inverse of digitizers matrix, and then left multiplied by the Kinects matrix. After that we could obtain the transformation matrix from digitizer coordinate system to Kinect coordinate system. With this transformation matrix, the position of digitized electrodes could convert their coordinate value into Kinect coordinate system.
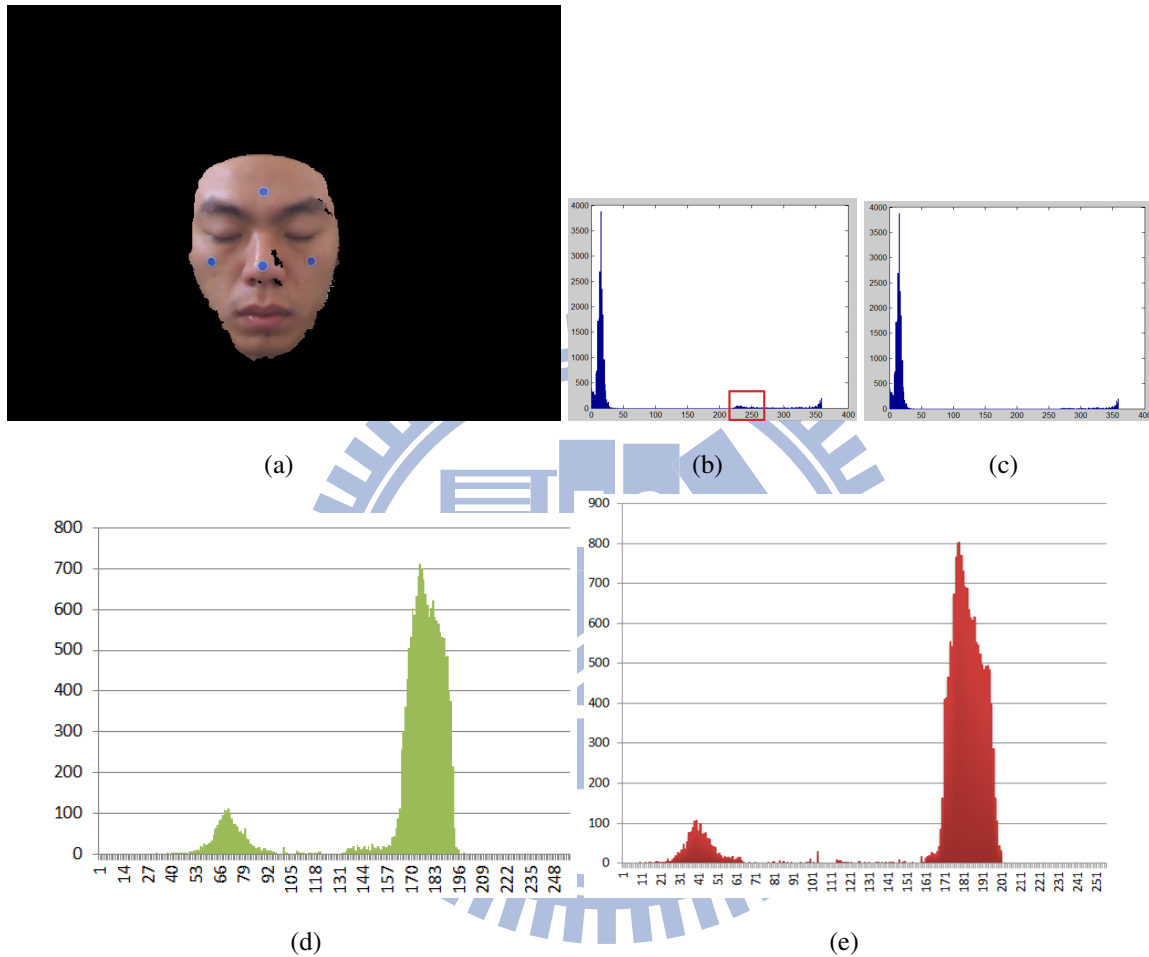
Figure 2.10: The analysis of facial surface colors included markers. (a) The facial surface. (b) The histogram of hue value of facial surface from (a), including the blue region. The blue region is in the red block. (c) The histogram of hue value of facial surface from (a), which is only skin region. (d) The histogram of red domain. (e) The histogram of green domain.

2.9   Thickness compensation

Because the electrodes on the EEG cap have a certain thickness, the electromagnetic digitizer digitizes the sensor position taped on the upper part of this support. To correct the electrode thickness so that digitized sensor coordinates lie on the scalp surface, the positions of the digitized target sensors points were adjusted. After coordinate transforms, all the point clouds (MRI-derived scalp points, Kinect-derived scalp points and digitized sensor points) in one coordinate system and were aligned together. For each digitized sensor points, found the closest MRI-derived scalp point and its neighbour points and calculated those points normal. To determine the correct sign for the normal, we calculate inner product between the vector from the centroid of MRI-derived scalp points to the closest MRI-derived scalp point and normal. If the result is positive, change the sign of normal, otherwise unchanged. By this way, we can adjust the direction of normal toward the inward direction from the scalp. After determine the correct sign for the normal, translate the digitized sensor point with a distance equal to the electrode thickness along the axis defined by the normal to obtain the scalp sensor point. Figure 2.12 is schematic diagram.

As for the electrode thickness, we measure all electrodes' thickness on the EEG cap by using localization device ten times. First, place the paper which draw the electrode-sized circle and its centroid on the table, and use localization device to record the centroid coordinate on the paper. Second, for each electrode on the EEG cap, aligned with the circle in the paper and record the electrode position coordinate. Third, calculate the distance between two positions coordinates and then average those distances as electrode thickness.

2.10   Error estimation

In order to evaluate the accuracy of registration and compared with different methods made before, we refer to Koreeler et al. [11] and used three criteria: one is residual error of face (REF) that is the average of Euclidean distance between each Kinect-derived facial points and the closest MRI-derived scalp point.

$$E_{REF} = \frac{1}{|a|} \sum_a \min_b \|a - b\| \qquad (2.8)$$

where E_{REF} is the residual error of face, $a \in$ Kinect-derived facial points and $b \in$ MRI-derived scalp points.

Another one is residual error of sensor (RES) which is the average of Euclidean distance between the scalp sensor points and the closest points of MRI-derived scalp points.

$$E_{RES} = \frac{1}{|a|} \sum_a \min_b \|a - b\| \tag{2.9}$$

where E_{RES} is the residual error of the sensor, $a \in$ scalp sensor points and $b \in$ MRI-derived scalp points.

In our case, the "true" error, or called map error which measure a distance between two real corresponding points does not exist because it requires the knowledge of which MRI point corresponds to which digitized point. And according to Whalen et al. [16], the target residual error (TRE), using the markers visible in MRI scans could replace this problem. Due to not accessible the TRE, we consult the residual error of sensor proposed by Koreeler et al. [11], as evaluation criteria that is more suitable for our method even though it underestimates the true registration errors. Besides, in order to check the error of transformation, we used digitization device to point the cross on the head in the digitizer coordination system, and after transformation we calculated the average of Euclidean distance between the cross points and the closest points of MRI-derived scalp points as three criteria called residual error of cross.

$$E_{REC} = \frac{1}{|a|} \sum_a \min_b \|a - b\| \tag{2.10}$$

where E_{REC} is the residual error of cross, $a \in$ digitized cross points and $b \in$ MRI-derived scalp points. That criterion does not consider the thickness compensation. Moreover, the intra- and inter- subject variability was calculated.

Figure 2.11: The result of estimation of centroid of markers. (a) The positions of markers on face. (b) After automatic estimation of centroid of markers, the centroids would found. The pink points are centroid of markers.
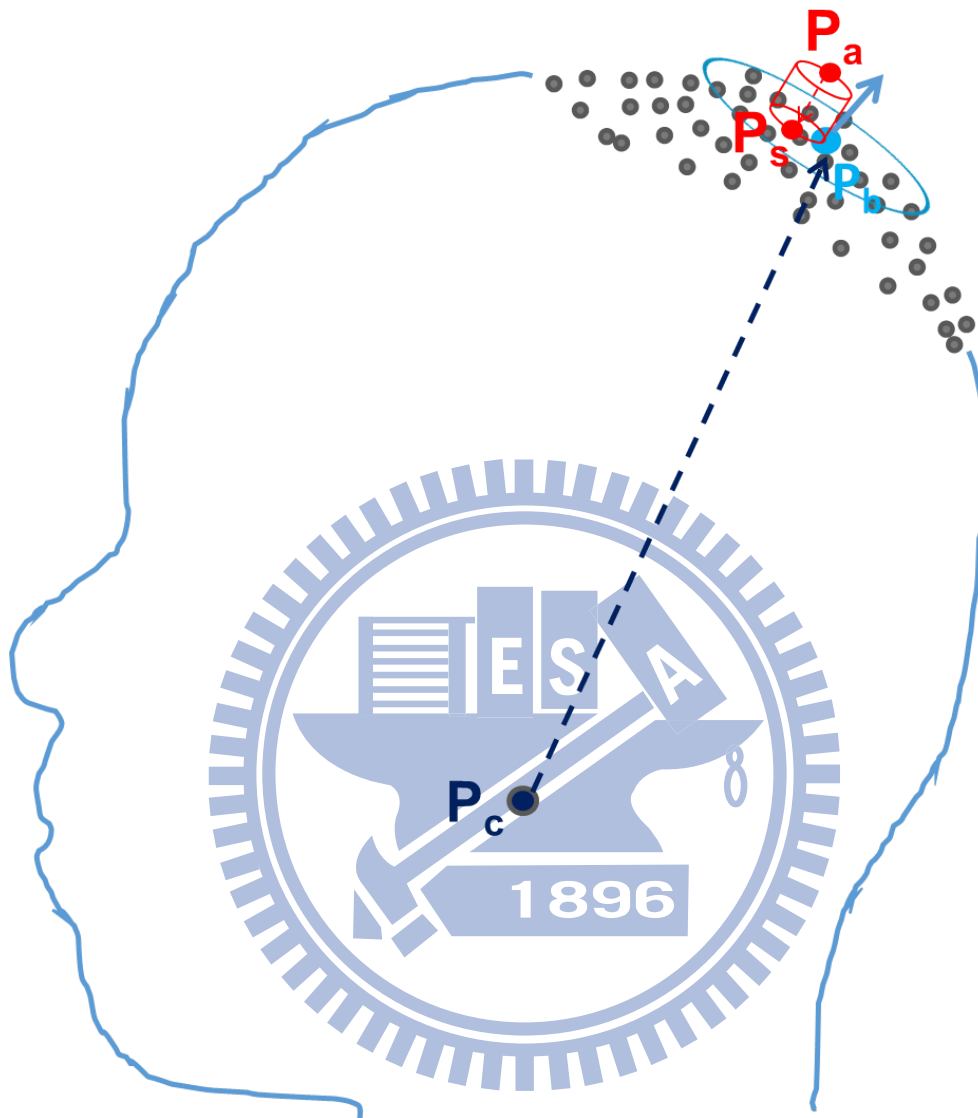
Figure 2.12: Schematic diagram of thickness compensation. The red point floating on the scalp surface is digitized sensor point P_a. At first, finding the closest MRI-derived scalp point P_b (blue point) and its neighbourhood (grey points in the blue circle) and calculated those points normal (blue arrow). To determine the correct sign for the normal, we calculated inner product between the vector (navy blue dotted line) from centroid of MRI-derived scalp points P_c to the closest MRI-derived scalp point P_b and normal. And translate the digitized sensor point with a distance equal to the electrode thickness along the axis defined by the normal to obtain the scalp sensor point P_s (red point that near scalp surface).

Chapter 3

Experimental Results

3.1   Experimental Setup

3.1.1   Subjects

Ten healthy subjects participated in this experiment, and do the experiment twice for each person. Before the experiment started, the face of subject would be pasted on four markers: one on the tip of nose, another on the middle of the forehead above eyebrows and the others are on the skin of right and left zygomatic bone. The reasons for choosing these places and the design of marker would be explained in the following section. Furthermore, the subject would wear a head-sized EEG cap which used for EEG recordings. In our case, The EEG cap contains 31 electrodes, and one of that is ground.

3.1.2   MRI

The MRI scans were acquired on Siemens MAGNETOM Trio, A Tim System 3T scanner (Siemens Medical Solution, Erlangen, Germany) with 12-Channel head coil. (See Figure 3.1.) We used a magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence (TR = 2530ms, TE = 3.03ms, TI = 1100ms, field of view = 224 x 256, matrix size = 224 x 256 and 192 continuous slices). The slice thickness is 1 mm, and voxel size is 1 x 1 x 1 mm.

Figure 3.1: Siemens MAGNETOM Trio, A Tim System 3T scanner.

3.1.3 Digitization

Any localization device that could acquire the locations of electrodes was able to be used in our experiment. Our study used a Polhemus FastTrak that contains two receivers and one transmitter. (See Figure 3.2) Two receivers include a stylus which points to the desired locations for collection and three receivers which are fixed to the head of subject to endure slight displacements of the head without producing digitization errors. The subject sits before the fixed transmitter which orientation with X pointing up and Y pointing to the right.

Figure 3.2: Polhemus FastTrak.

3.1.4 Experimental environment

We setup a pair of studio lamp beside the subject because of getting uniform color images and reducing the effect of the shadow. The Kinect for windows mounted in front of the subject and toward the subject, and subject near Kinect for windows as close as possible. The proposed system is implemented in C/C++ language, compiled by Microsoft Visual Studio 2010 compiler. OpenCV [3], OpenNI [4] and PCL [5] library are also applied in our system.

3.1.5 Blue marker

Before the experiment started, the subject would be pasted some markers that used to find the relation of transformation between Kinect and digitizer coordinate systems. In order to extract the markers conveniently, we choose a specially crafted marker that was 11 mm blue round, white outer ring and white cross line in the central. There were two reasons for choosing this design of makers: (i) we wanted the color be great different to

our skin so that we could easily find the range of marker from color information. (ii) To make sure the cross line would be in the central in the limited technology of printing and funding budget, the white outer ring and cross line were be printed in the meantime that is why we must have the white outer ring in the markers. Figure 3.3 shows the blue marker.

In our experiment, we pasted four markers on the face, and in the discussion section would give information that how many markers and where the appropriate positions of marker could bring the correct transformation relation.

Figure 3.3: Blue marker.

3.2 Results

Our study used three error estimations to evaluate the accuracy of co-registration, and we calculated them in root mean square (R.M.S) form and arithmetic mean form respectively. Figure 3.4 shows the Kinect-derived facial points and MRI-derived facial points in the same coordinate system before co-registration, and Figure 3.5 displays the result of co-registration of Kinect-derived facial points and MRI-derived facial points in the form of point cloud. Then Figure 3.6 and 3.7(a)-(i) shows example of Kinect-derived facial points superimposed on the different slices after co-registration. In this example, we could observe some noise and outlier that extracted by Kinect for windows floating in the air. Table 3.1 shows the result of the residual error of the face. The mean of residual error of the face was equal to 0.89 mm $\pm$ 0.09 mm (Mean $\pm$ SD). Table 3.2 shows our method compared with other methods made before. It demonstrates that our method has better results by

using facial points to do registration.

Table 3.1: The results of the residual error of the face. (unit: mm)

	R.M.S	Mean	Std. dev.	intra-subject
S1	1.00	0.90	0.58	0.02
	1.01	0.86	0.53	
S2	1.00	0.87	0.54	0.045
	0.99	0.78	0.50	
S3	0.94	0.80	0.52	0.06
	1.00	0.92	0.53	
S4	1.09	0.97	0.64	0.11
	1.06	0.75	0.54	
S5	1.13	1.02	0.61	0.02
	1.22	1.06	0.63	
S6	1.05	0.87	0.57	0.025
	1.07	0.92	0.65	
S7	0.99	0.82	0.56	0.00
	0.98	0.82	0.52	
S8	1.07	0.85	0.74	0.025
	1.09	0.90	0.74	
S9	1.01	0.85	0.53	0.015
	0.99	0.82	0.54	
S10	1.34	1.15	0.74	0.065
	1.13	1.02	0.64	

The results of the residual error of the sensor were listed in Table 3.3. The mean of the residual error of the sensor was equal to 1.67 mm $\pm$ 0.46 mm (Mean $\pm$ SD). The max value of the residual error of the sensor by different subjects were about 2.67 - 4.76 mm due to the hair that cause the digitized points floating on the head and the sized difference between the EEG cap and the head. In addition, The thickness compensation also affected the residual error of the sensor. On the other hand, we compared with Koessler et al. [11]

(a)

(b)

Figure 3.4: The Kinect-derived facial points and MRI-derived facial points (red points). There are put in the same coordinate system. (a) and (b) are two different perspectives.
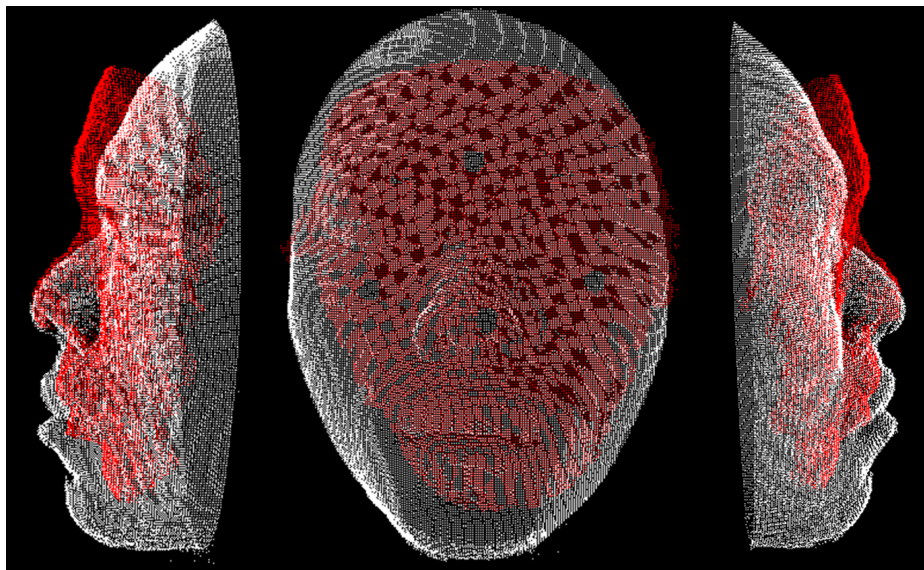
(a)
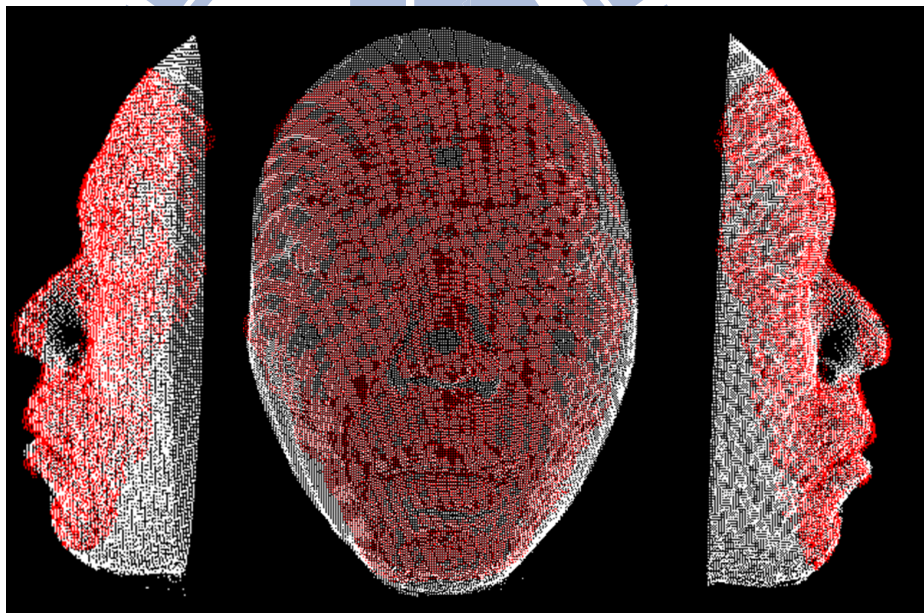
(b)

Figure 3.5: The result of co-registration of Kinect-derived facial points and MRI-derived facial points. (a) Before co-registration. The centroids of Kinect-derived facial points (white points) and MRI-derived facial points (red points) were matched together by translation. There are different perspectives. (b) After co-registration. The Kinect-derived facial points(white points) superimposed on the MRI-derived facial points (red points).

(a)

(b)

(c)
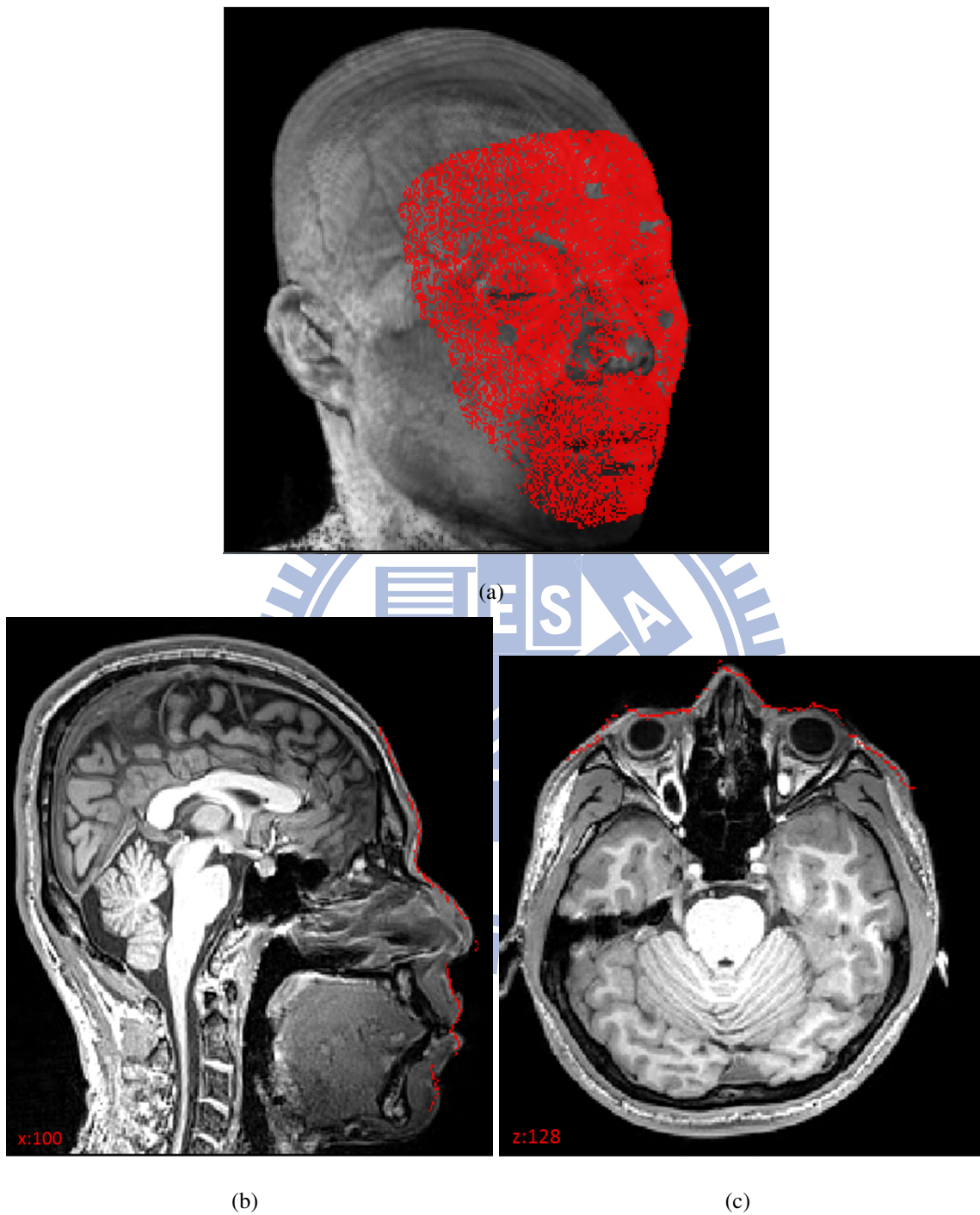
Figure 3.6: Example of the Kinect-derived facial points superimposed on MRI after co-registration. The red points are the Kinect-derived facial points. (a) The view from the right side. (b) and (c) are one of slices in vertical and horizontal tangent direction.
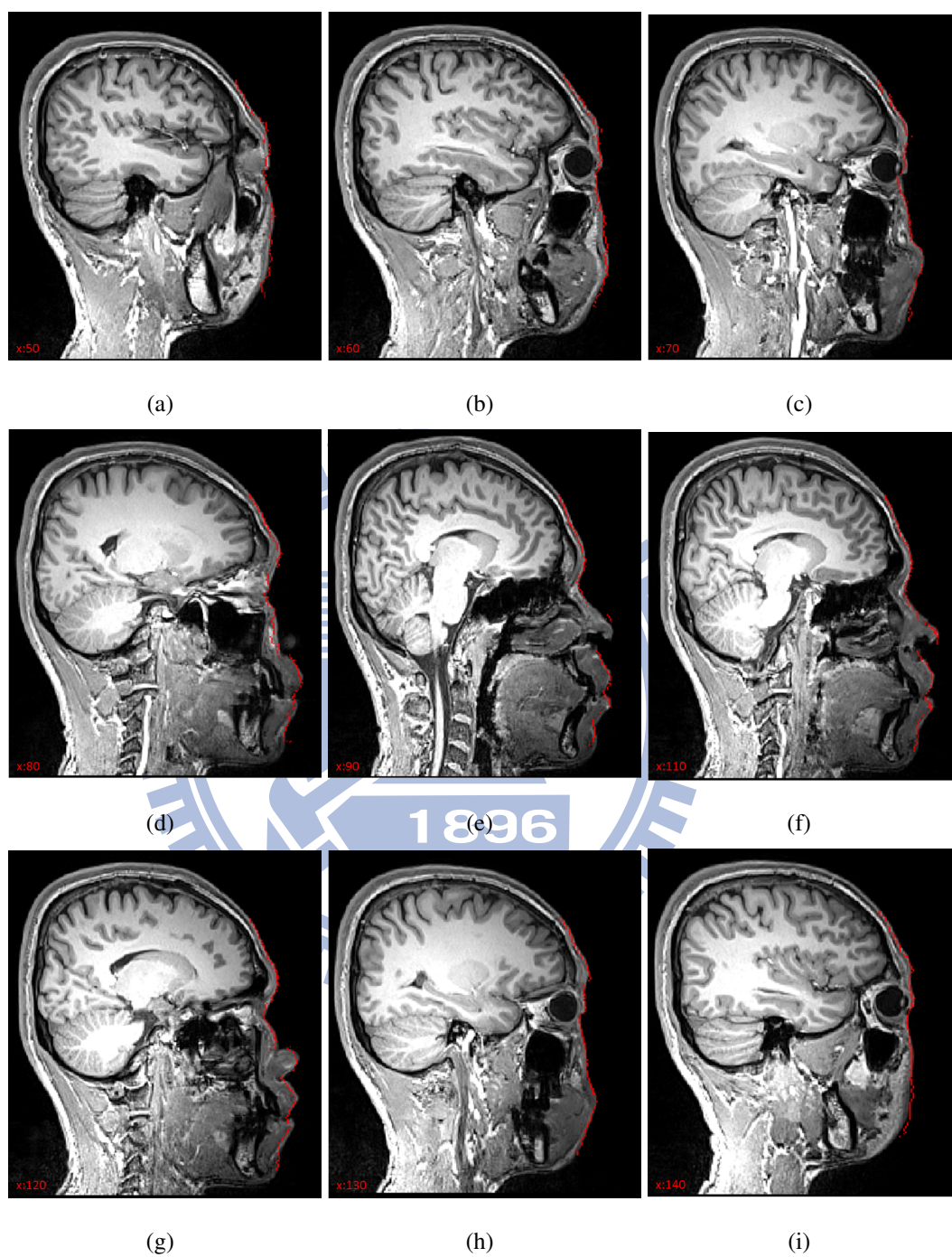
Figure 3.7: Example of Kinect-derived facial points on the MR slice. The red points are Kinect-derived facial points. (a) to (h) are different slices in vertical tangent direction.

Table 3.2: Comparison of the residual error of the face with other methods. (unit: mm)

	Our method	**Koessler et al.**	**Whalen et al.**
Subject	10 x 2	4 x 2	6
Mean	0.89 ± 0.09 Max = 1.15 Min = 0.75	1.68 ± 0.55 Max = 2.65 Min = 1.18	1.6 ± 0.1
R.M.S	1.05 ± 0.09 Max = 1.34 Min = 0.94	1.93 ± 0.91 Max = 3.5 Max = 1.19	-

that listed in Table 3.4, and our method had smaller errors. Figure 3.8 shows the example of digitized sensor points co-registered with MRI head surface, and Figure 3.8(c) can see the four makers on the face. Moreover, Figure 3.8(d) shows the sensors and the name of sensors, and Figure 3.9 shows example of digitized sensor locations on the MR slices.

By residual error of the cross, we could evaluate our system that use Kinect for windows as interface whether produce the great transformation errors. The results of the residual error of the cross were listed in Table 3.5. The mean of the residual error of the cross was equal to 1.83 mm $\pm$ 0.43 mm (Mean $\pm$ SD), and in R.M.S form was equal to 2.29 mm $\pm$ 0.48 mm. Then Figure 3.10 shows example of digitized cross locations on the head surface, and Figure 3.11 shows the example of digitized cross locations on the different MRI slices.

From ten subjects, we could observe that s9 and s10 have relative larger errors that because of our EEG cap only having medium and big size, the head of subjects s9 and subjects s10 were not fit the medium or big sized EEG cap. Simultaneously, the hair of subjects would influence the locations of digitized sensors. The hair of subject s5 was short and the head was medium size, so the results were relatively small.

Table 3.3: The results of the residual error of the sensor. (unit: mm)

	R.M.S	Mean	Std. dev.	Max	Min	intra-subject
S1	2.00	1.48	1.34	4.63	0.08	0.025
	1.76	1.43	1.01	3.84	0.21	
S2	2.00	1.14	1.64	5.84	0.05	0.21
	2.31	1.56	2.15	4.87	0.21	
S3	1.57	1.27	0.86	3.01	0.17	0.025
	1.68	1.32	1.12	2.67	0.13	
S4	2.11	1.93	0.86	3.87	0.11	0.105
	2.18	1.72	2.31	4.76	0.88	
S5	2.56	0.92	2.39	3.43	0.10	0.065
	2.17	1.05	3.14	4.33	0.17	
S6	2.4	1.97	3.41	2.43	0.03	0.03
	2.07	1.91	1.74	3.33	0.12	
S7	3.03	2.54	1.82	4.49	0.28	0.77
	1.57	1.00	1.21	3.23	0.12	
S8	2.78	1.93	1.34	3.49	0.15	0.205
	1.83	1.52	1.44	3.23	0.21	
S9	3.2	2.34	2.35	3.49	0.16	0.165
	3.17	2.01	2.27	3.34	0.15	
S10	3.47	2.31	3.51	5.47	0.35	0.075
	3.76	2.16	3.64	5.58	0.12	
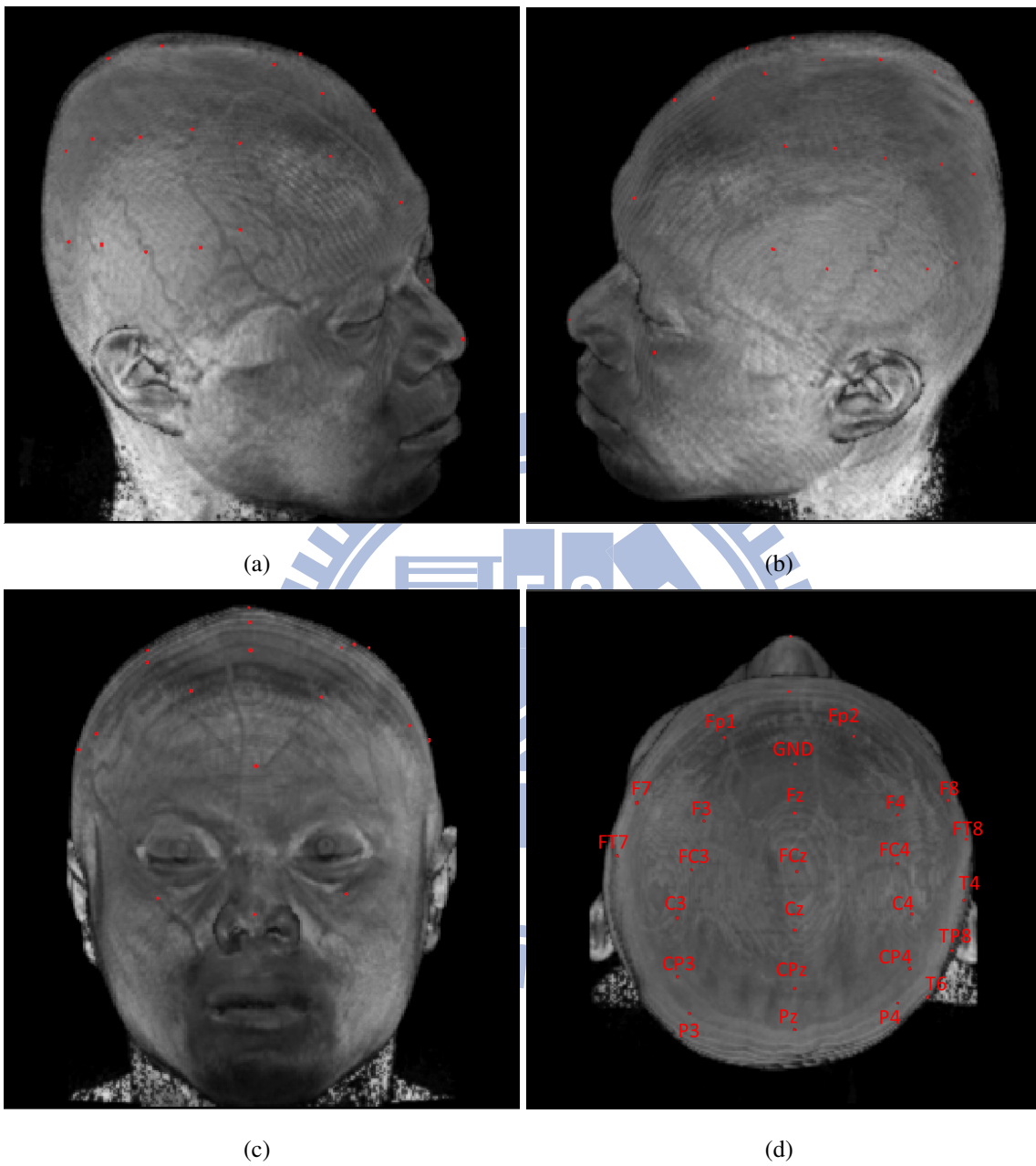
Figure 3.8: Example of digitized sensor points co-registered with MRI head surface. The red points are the locations of EEG sensors. (a) The right side of head surface. (b) The left side of head surface. (c) The front side of head surface. There are four markers on the face. (d)Looking down form height. The name of sensors were labelled.
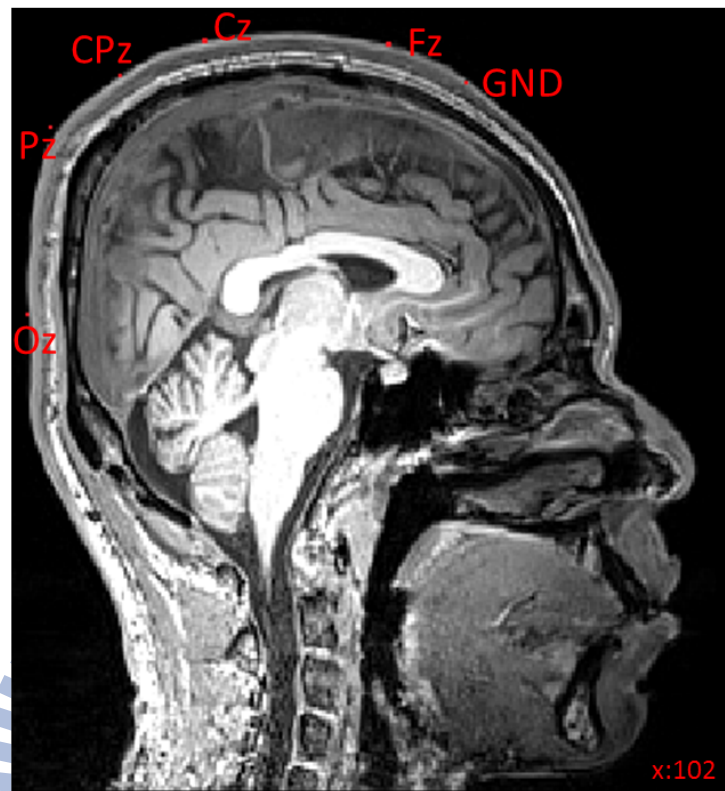
(a)

(b)

Figure 3.9: Example of digitized sensor locations on the MR slice. The red points are the locations of EEG sensors.

Table 3.4: Comparison of the residual error of the sensor with Koessker et al.. (unit: mm)

	Our method	**Koessler et al.**
Subject	10 x 2	4 x 2
Mean	1.67 ± 0.46	2.11 ± 0.46
	Max = 2.54	Max = 2.75
	Min = 0.92	Max = 1.62
R.M.S	2.38 ± 0.63	2.52 ± 0.63
	Max = 3.76	Max = 3.49
	Min = 1.57	Min = 1.80

(a)   (b)
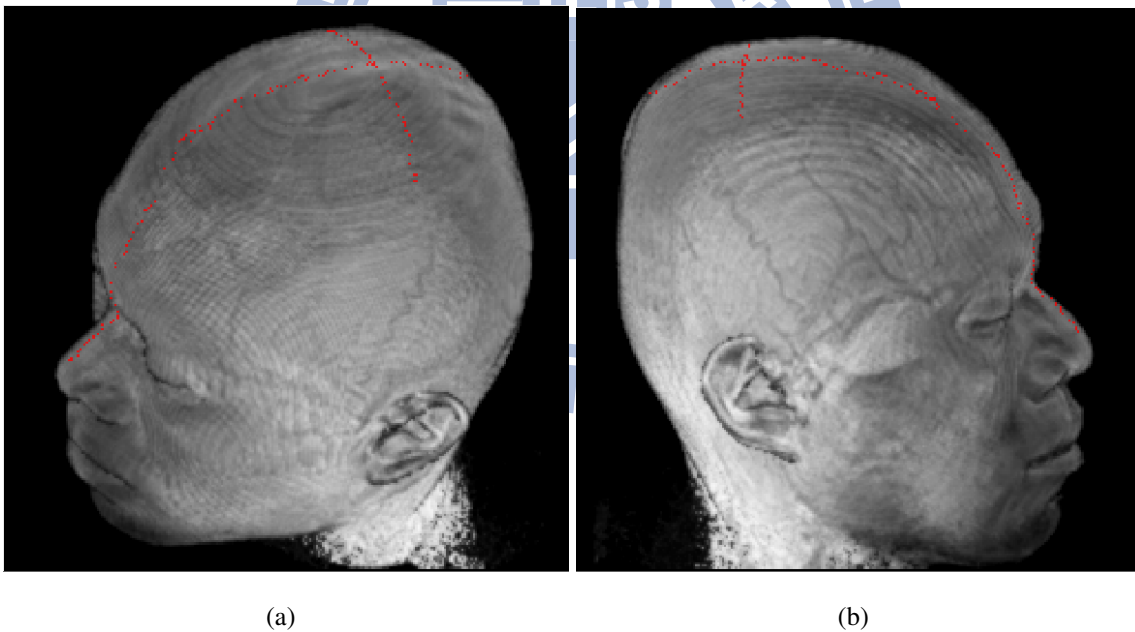
Figure 3.10: Example of digitized cross locations on the head surface. The red points are the digitized cross points. (a) The right side of head surface. (b) The left side of head surface.

Table 3.5: The results of the residual error of the cross. (unit: mm)

	Points	R.M.S	Mean	Std. dev.	intra-subject
S1	192	2.2	1.57	1.64	0.04
	208	1.93	1.49	1.37	
S2	152	2.34	2.01	1.28	0.07
	113	1.91	1.87	1.33	
S3	132	1.76	1.38	1.22	0.1
	127	1.97	1.58	1.37	
S4	134	2.34	2.31	2.17	0.035
	127	1.83	1.62	1.74	
S5	162	1.54	1.15	1.01	0.345
	173	2.07	1.32	1.59	
S6	158	2.33	1.83	2.17	0.165
	127	1.65	1.38	0.90	
S7	181	1.78	1.47	1.36	0.225
	170	2.56	2.09	1.48	
S8	212	2.42	1.97	1.41	0.478
	179	2.07	1.34	1.64	
S9	164	2.48	2.2	1.61	0.01
	178	2.79	2.22	1.69	
S10	157	3.07	2.78	1.64	0.16
	161	3.24	2.46	1.58	

(a)

(b)

(c)

Figure 3.11: Example of digitized cross locations on the MRI slices. The red points are the digitized cross points. (a), (b) and (c) are digitized cross locations on different MR slices.

Chapter 4

Discussions

4.1 Comparison of Kinect and Kinect for windows

In our study, we used Kinect for windows to extract the color and depth information. Compared with Kinect, Kinect for windows provides the near mode that not only extracts more points but also more precise in the same objects. The more precise these points were, the more accuracy of co-registration was. In this section we used two different RGB-D cameras to do registration between Kinect and MRI coordinate systems, and used the residual error of face to evaluate the accuracy of registration. Table 4.1 presents the average results by two different RGB-D cameras. The average result of registration using Kinect for windows was equal to 0.89 mm ± 0.03 mm (mean ± SD), and using Kinect was equal to 1.24 mm ± 0.07 mm. This experiment showed under the same subjects, using Kinect for windows to extract the RGB-D data to do registration had better results. The results of using Kinect to do registration can refer to Table 4.2.

Table 4.1: Comparison of Kinect and Kinect for windows. (unit: mm)

	Kinect for windows	**Kinect**
Mean	0.89 ± 0.03	1.24 ± 0.07
	Max = 0.95	Max = 1.31
	Min = 0.83	Min = 1.13

Table 4.2: The results of using Kinect to do registration. (unit: mm)

	Case 1	**Case 2**	**Case 3**	**Case 4**
Mean	1.23	1.31	1.3	1.13
Standard deviation	0.61	0.67	0.63	0.67
Maximum	3.04	3.3	2.97	3.41

4.2 The influence of co-registration by down-sampling

In order to speed up the registration between Kinect and MRI coordinate systems using facial points, we reduced the Kinect-derived facial points uniformly as GMM centroids to

do registration. But using down-sampling points may influence the results of registration. Therefore, we used the different degrees of down-sampling Kinect-derived facial points to align with MRI-derived facial points, and observed whether influence the results of registration and how much time cost. Table 4.3 represents case about this discussion. It shows that the larger degree of down-sampling is, the lower time is. Using down-sampling points almost does not affect the results of registration, and greatly enhances the speed.

Table 4.3: Example of the influence of co-registration by down-sampling. (unit: mm)

	Points	Time cost	RMS	Mean	Std dev.
Kinect for windows	17,392	1'44'38	1	0.86	0.5
Downsampling					
$1/2^2$	4,327	'25'28	1	0.86	0.51
$1/3^2$	1,927	'11'33	1	0.86	0.51
$1/4^2$	1,082	'06'53	1.01	0.86	0.52
$1/5^2$	683	'04'29	1.04	0.86	0.59

4.3 Estimation of transformation relation

In our study, these positions of markers were used to align with Kinect and digitizer coordinate systems. However, it was unknown that how many markers should be used and where the positions of markers should be pasted on to bring well transformation relation between Kinect and digitizer coordinate systems. Therefore, we provided some information about using of the markers in this section.

In this experiment, the subject was pasted on with nine markers. Figure 4.1(a) shows these positions of markers. The criteria about choosing the position of marker includes: (i) Flat area to use the color information of Kinect for windows to extract the entire round of markers and to calculate its centroids. (ii) Positions where were able to express the ups and downs and characteristics of human face. Our error estimation was defined as averaged these correspondence distances between these coordinate values of nine centroids of markers in the Kinect coordinate system and nine centroids of markers transfer from digitizer

to Kinect coordinate system by using transformation matrix. And the transformation matrix calculated with different number and positions of markers. Equation 4.1 presents our error estimation. The smaller the value was, the smaller the error between two coordinate systems and more precise.

$$\mathbf{E}_\theta = \frac{1}{N} \sum_{i=1}^{N} \left(\|\mathbf{x}_i - \mathbf{T}_\theta \mathbf{p}_i\|_2 \right),$$

$$\text{where } \mathbf{T}_\theta = \arg \min_{\mathbf{T}} \left(\sum_{q \in \theta} \left(\|\mathbf{x}_q - \mathbf{T}\mathbf{p}_q\|^2 \right) \right)$$

(4.1)

There are three subjects in this experiment, and everyone received twice. Figure 4.1(a) shows one of these cases. The horizontal axis represents using various numbers of markers, and the vertical axis indicates the value of error. The data were values of error estimation calculated by different transformation matrices, and that estimated by different numbers and positions of markers. For example, we chose four from nine markers and using every combination to calculate transformation matrix to evaluate error estimation, and the value of error would be showed on the axis of using four markers.

In Figure 4.1(b), we could observe a case of the errors that calculated by using different number of points to estimate transformation matrix. With using number of points increases, the selection of combinations of different points had little effect on the error. For example, seeing Figure 4.1(b) presents using different number points of median, Q1 and Q3. We could find different error for each other little difference by selecting seven different points to produce, and selecting four different points produce the errors disparity is quite large.

However, we found that using different number of points to calculate transformation matrix and evaluating the error estimation (for example: Choosing i number of points to do transformation matrix, for i=4 to 9), these errors we got from some combinations that would reach similar performance (the minimal error of registration), and even less than the errors we got from choosing more points. For example, the minimal of error in using four points was 0.69 mm which is the smallest in case 1. Besides, we integrated six data and it showed these features, seeing Figure 4.2. As a result, we were interested in the condition of using few points, which points were able to get the less error. Seeing Table 4.4, which shows the combinations of which points that get the minimal error by using

different number of points and median of error in using different number of points in case 1.

Table 4.4: In case 1, the combinations of which points that get the minimal error by using different number of points and median of error in using different number of points.

Points #	Minimal of error (mm)	Used points	Median of error (mm)
9	0.76	all	0.76
8	0.74	1 2 4 5 6 7 8 9	0.79
7	0.74	1 2 4 6 7 8 9	0.84
6	0.72	2 4 5 7 8 9	0.97
5	0.71	2 4 7 8 9	1.21
4	0.69	2 4 8 9	1.97

In the all combinations of four points, we selected the minimum error that were within fifteen percent (and the errors less than 1 mm) and polled these combinations to find those four points that occurs most frequently. (See Figure 4.3) Among those points, we found point #9, #8 and #4 having the higher frequency of appearance (higher than 0.4 of each case), and point #1, #2, #3 also performs well. However, considering the convenience for operating, the symmetry and facial feature of our face, we chose point #2. In order to more understanding which combination of four points could get the less error, we listed these top 15 smaller case errors of each six case. We found that the combination of (#2, #4, #8, #9) were top 3 in each case (except the case3 were second and third in case3_1, case3_2, the other two cases ranked first.) See Table 4.5.

For the purpose of verifying the fourth point accurately, we added a point to forehead. That is there were 10 markers on the face, and Figure 4.4 shows the positions of markers on face. We used the same method to estimate the transformation relation of different numbers and positions. Figure 4.5 shows the data of integrating four subjects which also owns the features of previous experiment. We also listed four data that were top 15 smaller error which calculated from using combinations of four points.(See Table 4.6) We found the combination of (#2, #5, #9, #10) and (#3, #5, #9, #10) were ranked in the top few, and positions of these points were the same as the previous experiment on the face. So we decided four positions on the face of subject in our study where one on the tip of

(a) (b)
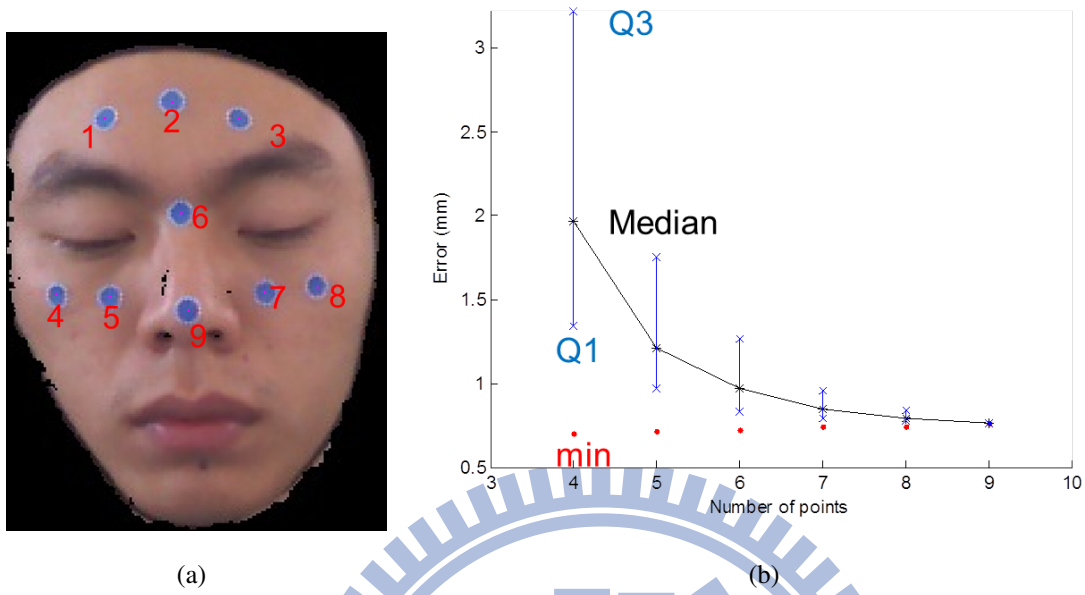
Figure 4.1: Example of estimation of transformation relation. (a) The positions of nine markers on face. (b)The error that calculated by using different number of points to estimate transformation matrix.
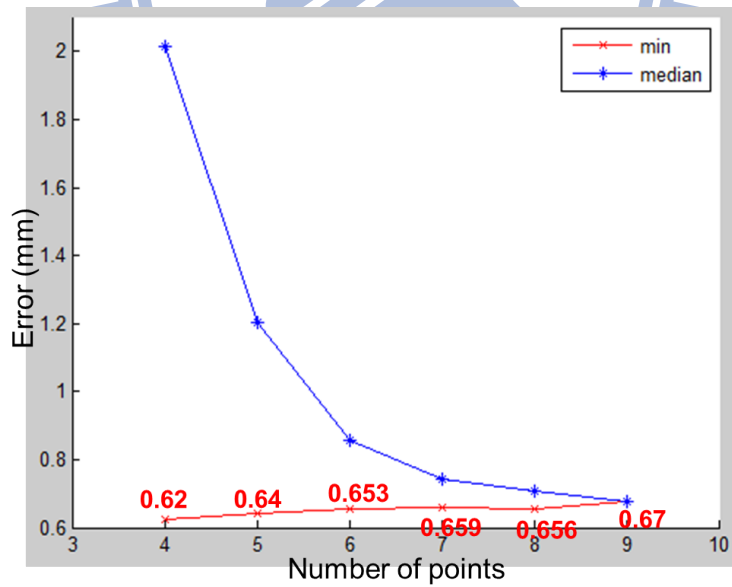

Figure 4.2: Average value of six data.
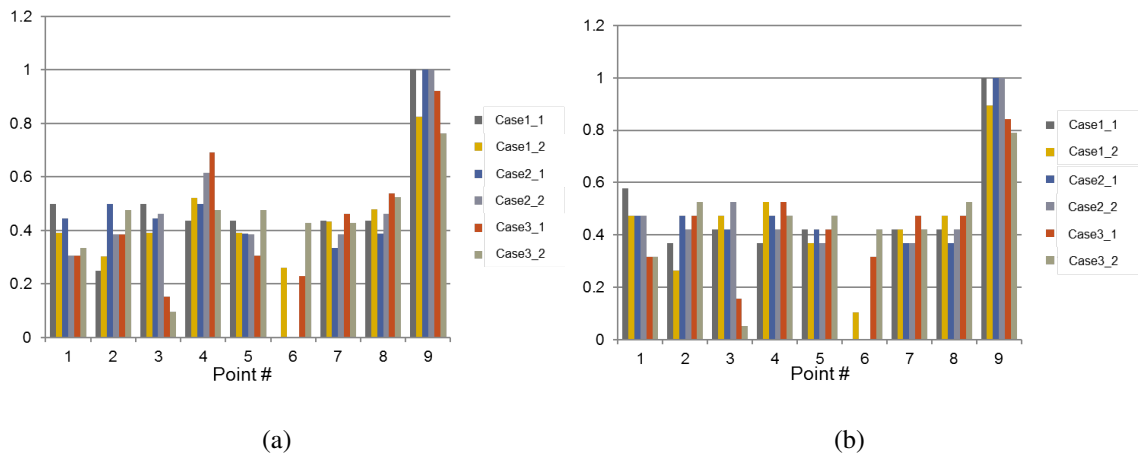
(a)   (b)

Figure 4.3: Frequency of occurrence of points. In the all combinations of four points, we selected the minimum error that were (b) within fifteen percent (a) and the errors less than 1 mm and polled these combinations to find those four points that occurs most frequently.

Table 4.5: The rank of smaller errors.

Rank	Case1_1	Case1_2	Case2_1	Case2_2	Case3_1	Case3_2
1	2 4 8 9	2 4 8 9	2 4 8 9	2 4 8 9	1 4 8 9	1 4 8 9
2	2 4 7 9	2 5 8 9	3 4 8 9	2 4 7 9	2 4 8 9	1 4 7 9
3	2 5 7 9	2 4 7 9	2 4 7 9	3 4 8 9	1 4 7 9	2 4 8 9
4	2 5 8 9	2 5 7 9	3 4 7 9	2 5 8 9	2 4 7 9	2 4 7 9
5	1 4 8 9	4 6 8 9	1 4 8 9	3 4 7 9	2 5 8 9	1 5 8 9
6	1 5 7 9	1 5 8 9	2 5 8 9	1 4 8 9	1 5 8 9	3 4 7 9
7	1 5 8 9	1 4 8 9	1 4 7 9	1 4 7 9	2 5 7 9	1 3 7 9
8	1 4 7 9	2 4 6 8	1 3 4 9	2 5 7 9	1 5 7 9	3 4 8 9
9	4 6 8 9	5 6 8 9	3 5 8 9	3 5 8 9	3 4 8 9	2 5 8 9
10	3 4 7 9	2 5 6 8	2 5 7 9	1 3 4 9	3 4 7 9	1 5 7 9
11	2 4 6 8	1 5 7 9	2 3 4 9	3 5 7 9	3 5 8 9	1 3 4 9
12	4 6 7 9	1 4 7 9	3 5 7 9	1 2 7 9	1 3 4 9	2 5 7 9
13	3 4 8 9	4 6 7 9	1 5 8 9	2 3 4 9	3 5 7 9	1 3 8 9
14	3 5 7 9	5 6 7 9	1 3 5 9	1 2 8 9	1 3 7 9	3 5 8 9
15	1 2 5 9	1 2 5 9	1 3 7 9	1 2 4 9	1 3 8 9	3 5 7 9

nose, another on the middle of the forehead above eyebrows and the others are on the skin of right and left zygomatic bone.
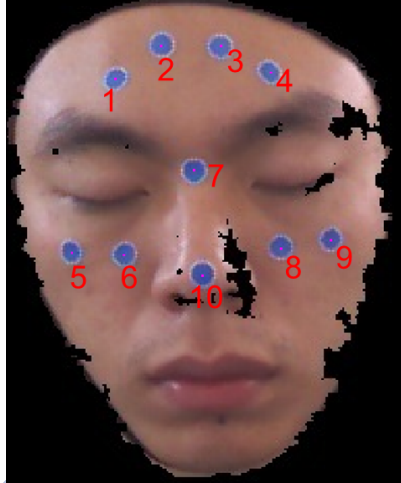

Figure 4.4: The positions of ten markers on face.

4.4 Source of error

The errors come from several sources, and we discussed the sources of errors divided into three error estimation individually. The influence of residual error of the face is (i) The error of points of coordinate values that extracted from Kinect for windows (including outliers and noise), (ii) scalp segmentation from MRI and (iii) the algorithm to do co-registration. The results showed that the residual error of the face relatively small. Concerning the influence of residual error of the sensor: (i) the sensor thickness and thickness compensation: different thickness compensation could generate different result due to different calculation of the normal. Certainly, the positions of sensors on the scalp would be different by different thickness compensation. (ii) EEG cap size and head of subject: the EEG cap must fit the head of subject, otherwise could influence the digitized coordinate values of sensors. (iii) Hair: The hair could bring the sensors on the EEG cap not fit to scalp of head, and influencing the digitized coordinate values. (iv) Camera resolution: Kinect for windows produces data stream about 640 x 480 and the resolution is lower compared

Table 4.6: Top 15 smaller error which calculated from using combinations of four points.

Rank	Case1_1	Case1_2	Case2	Case3
1	3 5 9 10	2 5 9 10	2 5 9 10	3 5 9 10
2	2 5 9 10	2 6 9 10	4 6 9 10	2 5 9 10
3	3 6 9 10	3 5 9 10	2 6 9 10	1 5 9 10
4	2 6 9 10	3 6 9 10	3 5 9 10	4 5 9 10
5	4 6 9 10	2 3 9 10	1 4 6 10	3 6 9 10
6	2 4 6 10	2 3 6 10	3 6 9 10	2 4 5 10
7	4 5 9 10	2 3 5 10	1 3 9 10	1 4 5 10
8	2 4 5 10	4 6 9 10	2 4 6 10	2 4 6 10
9	1 4 6 10	4 5 9 10	1 4 9 10	4 6 9 10
10	2 3 6 10	3 5 8 10	1 6 9 10	2 6 9 10
11	1 4 5 10	3 6 8 10	1 5 9 10	1 3 5 10
12	3 5 8 10	2 4 6 10	1 4 8 10	1 4 9 10
13	2 5 8 10	2 5 8 10	3 6 8 10	1 3 9 10
14	1 6 9 10	2 4 5 10	2 5 8 10	3 4 5 10
15	1 5 9 10	2 6 8 10	3 5 8 10	1 4 6 10

with current camera. (v) Digitizer error. (vi) Transformation relation: the error resulting from transformation relation which calculated by corresponding points and estimation of markers of centroids in Kinect coordinate system. As to residual error of the cross, there are no need to consider the sensor thickness and thickness compensation, but having some errors that come from (i) EEG cap size and head of subject; (ii) hair; (iii) digitizer error; (iv) transformation relation.
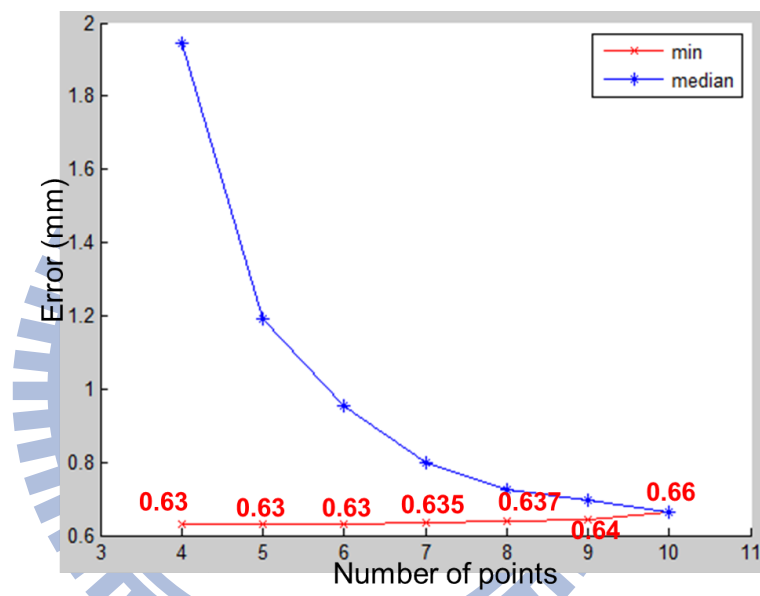
Figure 4.5: Average value of four data.

Chapter 5

Conclusion

In our study, we developed an automatic system to co-register EEG and MRI data used for the source estimation of electric brain activity. Our study used Kinect for windows that could extract the depth and color information as interface to align with MRI and digitizer coordinate systems. At first, we extracted points on facial surface from MR images that aligned with facial points which extracted by Kinect for windows. After reducing points uniformly and initial translation, we modified the coherent point drift algorithm to do registration between Kinect and MRI coordinate systems. Moreover, the purpose of modified coherent point drift algorithm and reduced points uniformly is to increase the speed and accuracy. Second, we estimated the centorids of markers that using color information provided by Kinect for windows and connected component automatically. By any kind of digitization device to record corresponding points on face, we used the coordinate correspondences to find the transformation relation between digitizer and Kinect coordinate systems. On the other hand, thickness compensation was applied to adjust the digitized points floating in the scalp surface. By the way, our work could be included any scalp-placed sensor co-registration.

According to the error estimation by Koessler et al. [11], the results of error estimation showed that our system was sufficiently accuracy. Our system was more accurate than Koessler et al. [11], and no need to additional labor and time to record facial feature. Besides, our system was less influence caused by variation of locating the position of the markers. Based on the above motioned, our system was sufficiently accurate, repeatable, efficient and labor-intensive to be used to assist neuroscience studies.

Bibliography

[1] Hsl color space. http://www.workwithcolor.com/hsl-color-schemer-01.htm.

[2] Kinect for windows blog. http://blogs.msdn.com/b/kinectforwindows/.

[3] opencv. http://opencv.org/.

[4] Openni tutorial. http://openni.org/docs2/Tutorial/index.htmll.

[5] Pcl tutorial. http://pointclouds.org/documentation/tutorials/.

[6] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Robotics-DL tentative*, pages 586–606. International Society for Optics and Photonics, 1992.

[7] Douglas Chai and King N Ngan. Face segmentation using skin-color map in videophone applications. *Circuits and Systems for Video Technology, IEEE Transactions on*, 9(4):551–564, 1999.

[8] Iok-Long Chan. 3-d object model reconstruction and immersive interaction using kinect sensor. Master's thesis, 2012.

[9] H-J Huppertz, M Otte, C Grimm, R Kristeva-Feige, T Mergner, and CH Lücking. Estimation of the accuracy of a surface matching technique for registration of eeg and mri data. *Electroencephalography and clinical neurophysiology*, 106(5):409–415, 1998.

[10] JR Ives, S Warach, F Schmitt, RR Edelman, and DL Schomer. Monitoring the patient's eeg during echo planar mri. *Electroencephalography and clinical neurophysiology*, 87(6):417–420, 1993.

[11] Laurent Koessler, Thierry Cecchin, Olivier Caspary, Adnane Benhadid, Hervé Vespignani, and Louis Maillard. Eeg–mri co-registration and sensor labeling using a 3d laser scanner. *Annals of biomedical engineering*, 39(3):983–995, 2011.

[12] D Kozinska, F Carducci, and K Nowinski. Automatic alignment of eeg/meg and mri data sets. *Clinical Neurophysiology*, 112(8):1553–1561, 2001.

[13] Andriy Myronenko and Xubo Song. Point set registration: coherent point drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(12):2262–2275, 2010.

[14] GV Simpson, ME Pflieger, JJ Foxe, SP Ahlfors, HG Vaughan, J Hrabe, RJ Ilmoniemi, and G Lantos. Dynamic neuroimaging of brain function. *Journal of Clinical Neurophysiology*, 12(5):432–449, 1995.

[15] Furlong P et al. Singh K, Holiday I. Evaluation of mri-eeg/meg coregistration strategies using monte carlo simulation. *Electroencephalography and clinical neurophysiology*, 102:81–85, 1997.

[16] Christopher Whalen, Edward L Maclin, Monica Fabiani, and Gabriele Gratton. Validation of a method for coregistering scalp recording locations with 3d structural mr images. *Human brain mapping*, 29(11):1288–1301, 2008.

[17] Seung-Schik Yoo, Charles RG Guttmann, John R Ives, Lawrence P Panych, Ron Kikinis, Donald L Schomer, and Ferenc A Jolesz. 3d localization of surface 10–20 eeg electrodes on high resolution anatomical mr images. *Electroencephalography and clinical neurophysiology*, 102(4):335–339, 1997.