most users. `SurfStat` was developed to utilize a model formula and avoids the explicit use of design matrices and contrasts, which tend to be a hinderance to most end users not familiar with such concepts. `SurfStat` can import MNI [237], FreeSurfer (`surfer.nmr.mgh.harvard.edu`) based cortical mesh formats as well as other volumetric image data. A similar model formula approach is implemented in many statistics packages such as Splus (`www.insightful.com`) R (`www.r-project.org`) and SAS (`www.sas.com`). These statistics packages accept a linear model like

$$\texttt{P} = \texttt{Group} + \texttt{Age} + \texttt{Brain}$$

as the direct input for linear modeling avoiding the need to explicitly state the design matrix. `P` is a $n \times 3$ matrix of coordinates, `Age` is the age of subjects, `Brain` is the total brain volume of subject and `Group` is the categorical group variable. This type of model formula has yet to be implemented in widely used SPM or AFNI packages. Here we illustrate `SurfStat` package by showing the step-by-step command lines for multivariate linear models.

7.5 Case Study: Surface Deformation Analysis

The data published in [206, 205] is used throughout the section. We have 3 Tesla T1-weighted inverse recovery fast gradient echo MRI, collected in 124 contiguous 1.2-mm axial slices (TE=1.8 ms; TR=8.9 ms; flip angle = 10°; FOV = 240 mm; 256 × 256 data acquisition matrix) of 69 middle aged and elderly adults ranging from 38 to 79 years (mean age = 58.04 ± 11.34 years). The data were originally collected for a national study on the health and well-being in the aging population, called MIDUS (midlife in US). The website explaining additional details of the study is given in `midus.wisc.edu`.

There are 23 males and 46 females in the study. Trained raters manually segmented the amygdala and hippocampus structures from the T1-weighted images. Brain tissues in the MRI scans were automatically segmented using Brain Extraction Tool (BET) [332]. Then we performed a nonlinear image registration using the diffeomorphic shape and intensity averaging technique with cross-correlation as similarity metric through Advanced Normalization Tools (ANTS) [19]. Using the deformation field obtained from warping the individual image to the template, we aligned the amygdala and hippocampus binary masks to the template space. The normalized masks were then averaged to produce the subcortical structure template. The isosurfaces of the subcortical structure template were extracted using the marching cube algorithm [232]. The amygdala and hippocampus surface template is given in Figure 7.6.

The whole data set is saved in a single `MAT` file `midus.mat` that contains the surface coordinates of the template, displacement vector fields for all 69 subjects, age and gender of the subjects. The following codes for

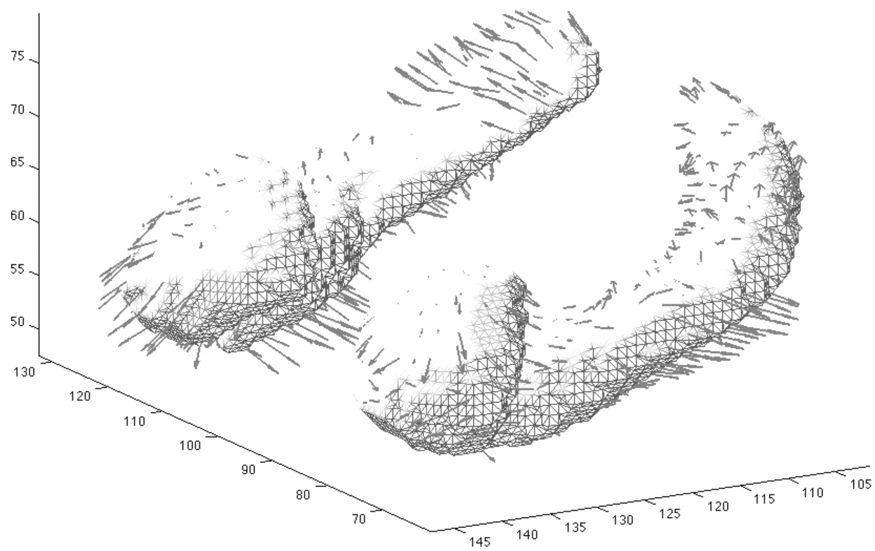

FIGURE 7.5
Displacement vector field of the 10th subject on top of the template surfaces. The end points of the arrows are where the surfaces of the 10th subject are.

performing univariate and multivariate general linear models are given in `http://brainimaging.waisman.wisc.edu/~chung/BIA`. Figure 7.6 is generated by running the codes:

```
load midus.mat

amygl=template{1,1}
amygr=template{1,2}
hippol=template{2,1}
hippor=template{2,2}

figure; figure_wire(amygl,'black','red')
hold on; figure_wire(amygr,'black','blue')
hold on; figure_wire(hippol,'black','yellow')
hold on; figure_wire(hippor,'black','green')
view([-120 15])
```

The left amygdala surface template is given as the structured array variable `amygl` and the right hippocampus template is given as `hippor`. Other surface variable names are similarly given. Each surface has a different number of mesh vertices and face numbers. For instance, `amygl` has the data structure of the form

```
amygl =

    vertices: [1296x3 double]
       faces: [2588x3 double]
```

The surface `amygl` has 1296 mesh vertices and 2588 triangles. At each mesh vertex, we have the displacement vector field. The displacement vector field is stored in the variable `displacement`, which is $69 \times 2 \times 2$ cells. Since there are 4 individual structures, there are 4 cell entries. The displacement vectors of the 10th subject left amygdala is accessed by `displacement{10,1,1}`. The cell entry is a usual matrix format, which is stored in the variable `dispal`. Other vector fields are stored similarly:

```
dispal = displacement{10,1,1};
dispar = displacement{10,1,2};
disphl = displacement{10,2,1};
disphr = displacement{10,2,2};
```

The variable `dispal` is a matrix of size 1296×3. The first few entries are

```
dispal =

   -2.4857    0.2103   -0.3367
   -2.3826    0.3100   -0.0352
   -2.9083    0.0469   -0.2653
   .
   .
   .
```

This is the response vector we will use in MGLM. Four vector fields `dispal`, `dispar`, `disphl` and `disphr` are displayed separately (Figure 7.5) as

```
figure;
hold on; figure_quiver2(amygl, dispal, 9)
hold on; figure_quiver2(amygr, dispar, 9)
hold on; figure_quiver2(hippol, disphl, 9)
hold on; figure_quiver2(hippor, disphr, 9)
view([-120 15])
```

Since the vector fields are dense, we only subsample 1 out of every 9 vectors and display. The subsampling rate is controlled in the last argument.

7.5.1 Univariate Tests in SurfStat

In the data set, we have `brain`, the total volume of brain excluding cerebellum in terms of mm^3, and `age` in terms of years. `gender` is the categorical variable consisting of 1 (female) and 0 (male).

```
gender =

     1     1     1     1     1     0  ...
```

SurfStat requires 3D array of vector fields to be arranged as 69 (number of subjects) $\times$ 2444 (number of vertices) $\times$ 3 (dimension). For example, the displacements for the left hippocampus surfaces `disphl` are given by

```
 disphl=[];
for i=1:69
    disphl(i,:,:)=displacement{i,2,1};
end

size(disphl)

ans =

          69        2444           3
```

From now on, we will only show statistical analysis for the left hippocampus since we are applying the same procedures to the other three surfaces as well. Let us test for the effect of gender on the length of displacement while accounting for brain size and age variations. This is done by the t-test via `SurfStatT` routine. The example below will be self-evident.

```
disphl2 = sqrt(sum(disphl.^2,3));

Gender=term(gender);
Age=term(age);
Brain=term(brain);

slm1 = SurfStatLinMod( disphl2, 1+ Age +Brain + Gender);
slm = SurfStatT( slm1, gender);
hold on; figure_trimesh( hippol ,slm.t)
```

Equivalently we can test for the effect of gender using the F-test via `SurfStatF`. The problem with the F test is that it does not provide the directional information so the T-test is recommended for testing one contrast (Figure 7.7).

```
slm0 = SurfStatLinMod( disphl2, 1+Brain +Age);
slm1 = SurfStatLinMod( disphl2, 1+Brain + Gender+Age);
slm = SurfStatF( slm1, slm0);
hold on; figure_trimesh( hippol ,slm.t)
```

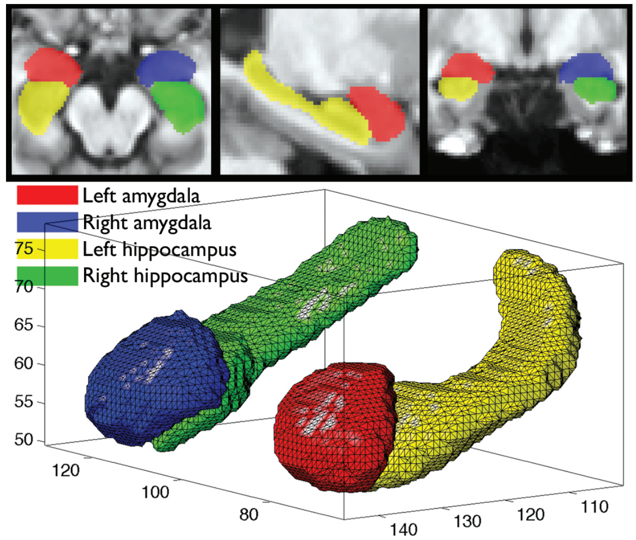

FIGURE 7.6
Subcortical structure template is superimposed on top of MRI [206, 205]. The figure is generated by Seung-Goo Kim of Seoul National University.

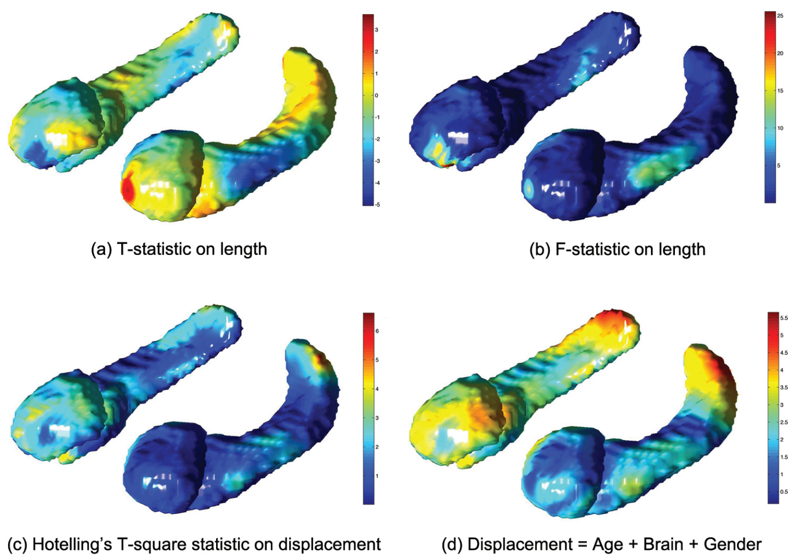

FIGURE 7.7
The gender effect on the length of displacement using (a) t-statistic and (b) F-statistic. The effect of gender on the displacement vector using (c) Hotelling's T^2 statistic and (d) Roy's maximum root.

7.5.2 Multivariate Tests in `SurfStat`

The problem with the univariate tests is that they are based on data reduction of the displacement vector field into a scalar field. So we are not utilizing full information for the data analysis. In order to use the vector field without the data reduction, we need MGLM. Let us test the effect of `gender` on the displacement vectors using the Hotelling's T^2 statistic, which is implemented in `hotelT2`. For this, we need to separate the displacement into male and female groups using `find` function as follows.

```
disphlm = disphl(find(~gender),:,:);
disphlf = disphl(find(gender),:,:);
h =SurfStat2HotelT2(disphlm, disphlf);
figure; figure_trimesh( hippol ,h.t)
```

Equivalently, we can take the maximum of the t-statistics for all possible linear combinations of the x, y and z coordinates such as $x + 0.5y - 1.2z$. This maximum t gives the square root of Hotelling's T^2. In `SurfStat`, statistical analysis proceeds exactly as for univariate data and uses the same statistical routines. The resulting Hotelling's T^2 value is displayed in Figure 7.7.

The problem with the Hotelling's T^2 statistic method is the lack of controls for age and brain size differences. This gives a motivation for using MGLM. To test the effect of `Gender`, the model of the from `displacement = Age + Brain + Group` is fitted using

```
slm0 = SurfStatLinMod( disphl, Age + Brain + Gender);
slm = SurfStatT( slm0, gender);
figure; figure_trimesh( hippol ,slm.t)
```

The result is given in Figure 7.7. In this chapter, we discussed the multivariate general linear models without surface-based smoothing and multiple comparisons issues, which will be covered in later chapters.