Regularized Discriminative Direction for Shape Difference Analysis

Luping Zhou[1], Richard Hartley[1,2], Lei Wang[1], Paulette Lieby[2], and Nick Barnes[2]

[1] RSISE, The Australian National University,
[2] Embedded Systems Theme, NICTA

Abstract. The "discriminative direction" has been proven useful to reveal the subtle difference between two anatomical shape classes. When a shape moves along this direction, its deformation will best manifest the class difference detected by a kernel classifier. However, we observe that such a direction cannot maintain a shape's "anatomical" correctness, introducing spurious difference. To overcome this drawback, we develop a *regularized* discriminative direction by requiring a shape to conform to its population distribution when it deforms along the discriminative direction. Instead of iterative optimization, an analytic solution is provided to directly work out this direction. Experimental study shows its superior performance in detecting and localizing the difference of hippocampal shapes for sex. The result is supported by other independent research in the same domain.

1 Introduction

It is critical to identify and understand the difference between two anatomical shape classes, such as normal/abnormal or male/female hippocampi. To differentiate linearly non-separable classes, kernel classifiers, such as SVMs, have been widely used. The class difference is identified in a high dimensional feature space $\mathcal{F}$ induced by a kernel function. While such difference is *mathematically* meaningful, it needs to be projected back to the shape descriptor space $\mathbb{R}^d$ to be explained in an *anatomically* meaningful way. For this purpose, Golland et al. proposed a "discriminative direction" method to isolate and visualize the subtle class difference in the shape descriptor space ([1, 2]). As in Figure 1, when a shape deforms along the discriminative direction towards the opposite class, the movement of its image in $\mathcal{F}$ will follow the direction $\mathbf{w}$ which best discriminates the two classes in $\mathcal{F}$. Hence the deformation can localize the class difference by ignoring the within-class variability. This method has been used to analyze the hippocampal shape difference between normal controls and patients [2]. In our previous work [3] we proposed a different way to calculate this discriminative direction and obtained the result similar to that of Golland's method.

* National ICT Australia is funded by the Australian Government's Backing Australia's Ability initiative, in part through the Australia Research Council. The authors thank the PATH research team at the Centre for Mental Health Research, ANU, Canberra, and the Neuroimaging Group (Neuropsychiatric Institute), Prince of Wales Hospital, Sydney, for providing the original MR and segmented data sets.

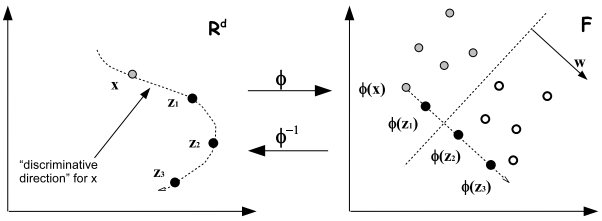


Fig. 1. *Discriminative direction for a point,* $\mathbf{x}$*. A nonlinear mapping* Φ *maps the shape descriptor space* $\mathbb{R}^d$ *onto a feature space* $\mathcal{F}$*, where the two classes (gray and white dots) become linearly separable. The* $\mathbf{w}$ *is the normal of a hyperplane found in* $\mathcal{F}$ *to best discriminate the two classes. Move* $\Phi(\mathbf{x})$ *along* $\mathbf{w}$ *to a new position* $\Phi(\mathbf{z}_1)$*, and project* $\Phi(\mathbf{z}_1)$ *back to* $\mathbb{R}^d$ *as* $\mathbf{z}_1$*. The vector* $\mathbf{z}_1 - \mathbf{x}$ *is the "discriminative direction" of* $\mathbf{x}$*. Similar results can be obtained for* $\Phi(\mathbf{z}_2)$ *and* $\Phi(\mathbf{z}_3)$*.*

However, we observed that simply deforming along this discriminative direction may introduce spurious shape differences. This is because the deformed shapes deviate from the underlying distribution of a shape class and untrue shapes are generated. That is, the intrinsic characteristic of a shape which makes it belong to a particular shape class (called "anatomical correctness" in this paper) is no longer maintained. In such cases, comparing the shapes before and after the deformation will lead to artifact differences, as shown later in the experiments. To remedy this, we argue that the shape distribution should not be ignored because (i) the shapes may only reside in a sub-dimensional manifold though the shape descriptor space has high dimensionality, and (ii) the deformation of an organ may be spatially restricted by its surroundings.

Our contributions are: (i) We identify the cause of the spurious difference as that deforming along the discriminative direction ignores the underlying shape distribution; (ii) We propose a regularized discriminative direction by requiring a shape to conform to the distribution when it deforms, and this is formulated as a penalized optimization problem; (iii) We derive an analytical solution for this optimization problem. It avoids performing optimization in a possibly high-dimensional shape descriptor space and directly works out the direction; (iv) After verifying our approach with controlled experiments, we analyze the shape difference of hippocampi for sex and compare it with that using the approach in [1]. We found our result agrees better with a published work about sex difference in hippocampal shapes [4] studied by a different approach.

2 Review of discriminative direction methods

Let $\mathcal{D} = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ ($\mathbf{x}_i \in \mathbb{R}^d$) be a set of n training samples from two classes. A kernel classifier implicitly performs a mapping $\Phi(\cdot)$ from the input space $\mathbb{R}^d$ to the feature space $\mathcal{F}$. An optimal separating hyperplane is obtained in $\mathcal{F}$ as $f(\mathbf{x}) = \mathbf{w}^\top \Phi(\mathbf{x}) + b$ where $\mathbf{w}$ is the normal and b is a bias. The vector $\mathbf{w}$ indicates the direction that best discriminates the two classes. Ideally $\Phi(\mathbf{x})$ should move

along $\mathbf{w}$ strictly to reflect only the class difference. However there is a dilemma. If $\Phi(\mathbf{x})$ moves strictly along $\mathbf{w}$, the resulting images in $\mathcal{F}$ might not have a pre-image in $\mathbb{R}^d$. On the other hand, by enforcing that the pre-image of $\Phi(\mathbf{x})$ does exist, $\Phi(\mathbf{x})$ cannot move strictly along $\mathbf{w}$. Two different solutions to this problem are given in [1, 2] and in our previous paper [3]. In [1, 2], Golland's method searched for the direction $d\mathbf{x}$ in $\mathbb{R}^d$ under the constraint $\|d\mathbf{x}\| = \epsilon$. When $\mathbf{x}$ moves along $d\mathbf{x}$, the divergence of the displacement $d\mathbf{z} = \Phi(\mathbf{x} + d\mathbf{x}) - \Phi(\mathbf{x})$ from $\mathbf{w}$ will be minimized in $\mathcal{F}$. Note that the constraint of $\|d\mathbf{x}\| = \epsilon$ allows $d\mathbf{x}$ to be searched identically along *all* directions in $\mathbb{R}^d$. Clearly the underlying distribution of $\mathbf{x}$ is not taken into account. From a different perspective, our previous work makes the movement of $\Phi(\mathbf{x})$ strictly follow $\mathbf{w}$, and approximates the corresponding pre-images in $\mathbb{R}^d$ by minimizing the reconstruction error $\rho(\mathbf{z}) = \| \Phi(\mathbf{x}) + s\mathbf{w} - \Phi(\mathbf{z}) \|^2$. Here $\mathbf{z} \in \mathbb{R}^d$ denotes the estimated pre-image and s denotes the step of movement in $\mathcal{F}$. Our previous work also allows $\mathbf{z}$ to freely move in all directions in $\mathbb{R}^d$.

3 Our approach - Regularized Discriminative Direction

The key idea of our new approach is that, when seeking for the pre-image of a point in $\mathcal{F}$, the possible solutions are restricted into a certain region rather than the whole $\mathbb{R}^d$ as in [1–3]. Without loss of generality, assume that $\mathbf{w}$ has been normalized as a unit vector. Let $\hat{\mathbf{x}}$ denote a particular shape to be deformed. Moving $\Phi(\hat{\mathbf{x}})$ along $\mathbf{w}$ in $\mathcal{F}$ for a step s arrives at a new position $\Phi(\hat{\mathbf{x}}) + s\mathbf{w}$. Let $\mathbf{z}$ be the pre-image of $\Phi(\hat{\mathbf{x}}) + s\mathbf{w}$, representing the new shape of $\hat{\mathbf{x}}$ after deformation. The estimation of $\mathbf{z}$ is elaborated below.

We argue that to ensure "anatomical correctness", the pre-image $\mathbf{z}$ should comply with the probability distribution of $\mathbf{x}$. For example, if $\mathbf{x}$ resides in a low-dimensional manifold, $\mathbf{z}$ should reside in it too. Let $\epsilon(\hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\| \leq \epsilon_0$ be a neighborhood of $\hat{\mathbf{x}}$ and $p(\mathbf{x} |\ \mathbf{x} \in \epsilon(\hat{\mathbf{x}}))$ be an empirical probability density function of $\mathbf{x}$ in $\epsilon(\hat{\mathbf{x}})$ estimated from the training shapes. We model $p(\mathbf{x} |\ \mathbf{x} \in \epsilon(\hat{\mathbf{x}}))$ as a normal distribution [1] with mean $\boldsymbol{\mu} = \hat{\mathbf{x}}$ and covariance matrix $\boldsymbol{\Sigma}$. For an RBF kernel $k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2)$, moving $\Phi(\mathbf{x})$ with a sufficiently small step s in $\mathcal{F}$ can always ensure that $\mathbf{z}$ stays in $\epsilon(\hat{\mathbf{x}})$. Hence we require that $p(\mathbf{z})$ should be large enough, or equally $(\mathbf{z} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu})$ be adequately small, provided that $\boldsymbol{\Sigma}$ has full rank. In this way, the optimal $\mathbf{z}$ is defined as

$$\mathbf{z}^* = \arg\min_{\mathbf{z} \in \mathbb{R}^d} \rho(\mathbf{z}) + 2\eta \cdot (\mathbf{z} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \tag{1}$$

where $\rho(\mathbf{z})$ is defined in Section 2, and η ($\eta \geq 0$) is the regularization parameter. When η is 0, this problem reduces to that in our previous work in [3]. According to our observation, our algorithm is insensitive to η in a reasonably large range.

Consider the case when the shapes reside in a sub-dimensional manifold, causing $\boldsymbol{\Sigma}$ to be rank-deficient. Decompose $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma} = \boldsymbol{\Gamma} \boldsymbol{\Lambda} \boldsymbol{\Gamma}^\top$, where each column

[1] Using a more complicated model may be dangerous in the sense that its parameters may not be reliably estimated because the number of training samples in $\epsilon(\hat{\mathbf{x}})$ is quite limited in practice.

of $\mathbf{\Gamma}$ is an eigenvector and $\mathbf{\Lambda}$ is diag$\{\lambda_1, \cdots, \lambda_k, 0, \cdots, 0\}$. The λ_i is the i-th positive eigenvalue and k is the rank of $\mathbf{\Sigma}$. An optimal solution $\mathbf{z}^\star$ should satisfy:

$$\mathbf{z}^\star \in \{\mathbf{z} | (\mathbf{\Gamma}^\top (\mathbf{z} - \boldsymbol{\mu}))_i = 0 \text{ for } i = k+1, \cdots, d\} = \{\mathbf{z} | \mathbf{z} = \boldsymbol{\mu} + \hat{\mathbf{\Gamma}} \hat{\mathbf{\Lambda}}^{\frac{1}{2}} \mathbf{u}\} \tag{2}$$

where $\mathbf{u} \in \mathbb{R}^k$, the $\hat{\mathbf{\Lambda}}$ is diag$\{\lambda_1, \cdots, \lambda_k\}$, and $\hat{\mathbf{\Gamma}}$ contains the corresponding eigenvectors. This optimization problem is explained as follows. Let $\mathcal{M}$ be the manifold where the shapes reside, and $T_{\boldsymbol{\mu}}(\mathcal{M})$ be a tangent plane of $\mathcal{M}$ at $\boldsymbol{\mu}$. This tangent plane is spanned by the eigenvectors in $\hat{\mathbf{\Gamma}}$. Since a manifold can be approximated locally by its tangent plane, the result in (2) can be thought of as confining the solution $\mathbf{z}$ to the manifold $\mathcal{M}$. Moreover noting that the shapes do not necessarily isometrically distribute in $T_{\boldsymbol{\mu}}(\mathcal{M})$, our regularized method naturally incorporates such distribution information via the $\hat{\mathbf{\Lambda}}^{\frac{1}{2}}$ in (2). This makes it achieve better performance than merely projecting $\mathbf{z}$ which maximizes $\rho(\mathbf{z})$ onto the tangent plane [2], as shown later in experiments. Finally, the problem in (1) can be simplified by optimizing $\mathbf{u}$ as:

$$\mathbf{u}^\star = \arg\min_{\mathbf{u} \in \mathbb{R}^k} \rho(\boldsymbol{\mu} + \hat{\mathbf{\Gamma}} \hat{\mathbf{\Lambda}}^{\frac{1}{2}} \mathbf{u}) + 2\eta \cdot \mathbf{u}^\top \mathbf{u} \tag{3}$$

and $\mathbf{z}^*$ is computed by (2). This greatly reduces the number of parameters to estimate compared with directly optimizing over $\mathbf{z}$. Iterative optimization methods can be used to estimate $\mathbf{u}$. However, when k is large, optimizing $\mathbf{u}$ is still cumbersome. Below we propose a new differential equation based solution so that for a given step s, first $\mathbf{u}^\star$ and then $\mathbf{z}^\star$ can be directly worked out.

3.1 An analytic solution to the pre-image $\mathbf{z}^\star$

The problem in (3) is equivalent to maximizing $\langle \Phi(\hat{\mathbf{x}}) + s\mathbf{w}, \Phi(\boldsymbol{\mu} + \hat{\mathbf{\Gamma}} \hat{\mathbf{\Lambda}}^{\frac{1}{2}} \mathbf{u}) \rangle - \eta \cdot \mathbf{u}^\top \mathbf{u}$ provided $\langle \Phi(\mathbf{x}), \Phi(\mathbf{x}) \rangle$ is a constant, such as for an RBF kernel. Noting that $\mathbf{w}$ lies in a space spanned by the training samples: $\mathbf{w} = \sum_i \alpha_i \Phi(\mathbf{x}_i)$, $\alpha_i \in \mathbb{R}$, we maximize the expression below. Note that k indicates the kernel function.

$$\begin{aligned} f(s, \mathbf{u}) &= \langle \Phi(\hat{\mathbf{x}}) + s\mathbf{w}, \Phi(\boldsymbol{\mu} + \hat{\mathbf{\Gamma}} \hat{\mathbf{\Lambda}}^{\frac{1}{2}} \mathbf{u}) \rangle - \eta \cdot \mathbf{u}^\top \mathbf{u} \\ &= k(\hat{\mathbf{x}}, \boldsymbol{\mu} + \hat{\mathbf{\Gamma}} \hat{\mathbf{\Lambda}}^{\frac{1}{2}} \mathbf{u}) + s \sum_i \alpha_i k(\mathbf{x}_i, \boldsymbol{\mu} + \hat{\mathbf{\Gamma}} \hat{\mathbf{\Lambda}}^{\frac{1}{2}} \mathbf{u}) - \eta \cdot \mathbf{u}^\top \mathbf{u} \\ &\triangleq g(\mathbf{u}) + s \cdot h(\mathbf{u}) - \eta \cdot l(\mathbf{u}). \end{aligned}$$

For each given s, there will be a $\mathbf{u}^*$ which maximizes $f(s, \mathbf{u})$. This optimization problem is not convex and has multiple local maxima. We propose an approach which does not directly solve the optimization over $\mathbf{u}$ for a given s. Instead it makes use of the fact that $(0, \mathbf{0})$ is a global maximum of $f(s, \mathbf{u})$ and traces the change of the global maximum with respect to s. As long as $\mathbf{u}^*(s)$ is continuous and differentiable, our solution remains the global or at least local maximum.

[2] This approach is called "tangent plane projection" later in our experiments.

The change of $\mathbf{u}^*$ with respect to s can be considered as a curve $\mathbf{u}^*(s)$ in $\mathbb{R}^k$ parametrized by s, passing through $(0, \mathbf{0})$. The curve can be traced out by computing its tangent $\frac{d\mathbf{u}^*}{ds}$. We approximate f by a second order Taylor expansion

$$\begin{aligned} f(s,\mathbf{u}) \approx & g(\mathbf{u}_0) + \mathbf{J}_g(\mathbf{u}-\mathbf{u}_0) + \frac{1}{2}(\mathbf{u}-\mathbf{u}_0)^\top \mathbf{H}_g(\mathbf{u}-\mathbf{u}_0) \\ & + sh(\mathbf{u}_0) + s\mathbf{J}_h(\mathbf{u}-\mathbf{u}_0) + s\frac{1}{2}(\mathbf{u}-\mathbf{u}_0)^\top \mathbf{H}_h(\mathbf{u}-\mathbf{u}_0) \\ & - \eta l(\mathbf{u}_0) - \eta \mathbf{J}_l(\mathbf{u}-\mathbf{u}_0) - \eta\frac{1}{2}(\mathbf{u}-\mathbf{u}_0)^\top \mathbf{H}_l(\mathbf{u}-\mathbf{u}_0), \end{aligned} \tag{4}$$

where $\mathbf{J}$ and $\mathbf{H}$ are the Jacobian and Hessian of the functions g, h and l with respect to $\mathbf{u}$, evaluated at $\mathbf{u}_0$. Here $\mathbf{u}_0$ maximizes $f(s,\mathbf{u})$ when $s = s_0$. The first order derivative of f with respect to $\mathbf{u}$ vanishes at $\mathbf{u}_0$ and other extrema $\mathbf{u}^*$. Since $\frac{\partial f}{\partial \mathbf{u}}\Big|_{\mathbf{u}_0} = 0$, we have $s_0\mathbf{J}_h = -\mathbf{J}_g + \eta\mathbf{J}_l$. Combining it in $\frac{\partial f}{\partial \mathbf{u}}\Big|_{\mathbf{u}^*} = 0$ gives

$$\frac{d\mathbf{u}^*}{ds}\Big|_{s=s_0} = -(\mathbf{H}_g + s_0\mathbf{H}_h - \eta\mathbf{H}_l)^{-1}\mathbf{J}_h. \tag{5}$$

The curve of $\mathbf{u}^*(s)$ can be therefore traced out by

$$\mathbf{u}^{*(t)} = \mathbf{u}^{*(t-1)} + \frac{d\mathbf{u}^{*(t-1)}}{ds}\Big|_{s=0}(s_t - s_{t-1}); \quad \mathbf{z}^{*(t)} = \boldsymbol{\mu}^{(t-1)} + \hat{\boldsymbol{\Gamma}}^{(t-1)}[\hat{\boldsymbol{\Lambda}}^{(t-1)}]^{\frac{1}{2}}\mathbf{u}^{*(t)}$$

where $\mathbf{u}^{*(0)} = \mathbf{0}$, $\hat{\boldsymbol{\Gamma}}^{(t-1)}$ and $\hat{\boldsymbol{\Lambda}}^{(t-1)}$ are estimated from $\mathbf{z}^{*(t-1)}$, and $\mathbf{z}^{*(0)} = \hat{\mathbf{x}}$. A four-stage Runge Kutta method is integrated to suppress the lower-order error terms of this ordinary differential equation. Our algorithm is summarized below.

■ **Algorithm 1**

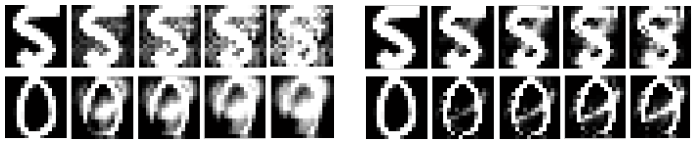
1. $\mathbf{z}_0 \equiv \hat{\mathbf{x}}, s_0 \equiv 0, \mathbf{u}_0 \equiv \mathbf{0}$
2. Estimate $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{\Lambda}}$ at $\mathbf{z}_0$; Evaluate $\mathbf{J}_h$, $\mathbf{H}_g$, $\mathbf{H}_h$ at $\mathbf{u}_0$ with an RBF kernel.
3. Compute $\mathbf{u}$ and the new position $\mathbf{z}$ using a four-stage Runge Kutta method.
   (1) $\mathbf{u}_1 \longleftarrow \mathbf{u}_0$, $s_{u1} \longleftarrow s_0$, compute $\mathbf{t}_1 = \frac{d\mathbf{u}^*}{ds}|_{s=s_{u1},\mathbf{u}=\mathbf{u}_0}$.
   (2) $\mathbf{u}_2 \longleftarrow \mathbf{u}_1 + \mathbf{t}_1\Delta s/2$, $s_{u2} \longleftarrow s_0 + \Delta s/2$, compute $\mathbf{t}_2 = \frac{d\mathbf{u}^*}{ds}|_{s=s_{u2},\mathbf{u}=u_0}$
   (3) $\mathbf{u}_3 \longleftarrow \mathbf{u}_1 + \mathbf{t}_2\Delta s/2$, $s_{u3} \longleftarrow s_0 + \Delta s/2$, compute $\mathbf{t}_3 = \frac{d\mathbf{u}^*}{ds}|_{s=s_{u3},\mathbf{u}=u_0}$
   (4) $\mathbf{u}_4 \longleftarrow \mathbf{u}_1 + \mathbf{t}_3\Delta s/2$, $s_{u4} \longleftarrow s_0 + \Delta s$, compute $\mathbf{t}_4 = \frac{d\mathbf{u}^*}{ds}|_{s=s_{u4},\mathbf{u}=u_0}$
   (5) Compute $\mathbf{u} = \mathbf{u}_0 + \Delta s \times (1/6\mathbf{t}_1 + 1/3\mathbf{t}_2 + 1/3\mathbf{t}_3 + 1/6\mathbf{t}_4)$
   (6) Compute the new position $\mathbf{z}$: $\mathbf{z} = \boldsymbol{\mu} + \hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{\Lambda}}^{\frac{1}{2}}\mathbf{u}_0$
4. $\hat{\mathbf{x}} \longleftarrow \mathbf{z}$
5. Repeat step 1 $\sim$ 4 to get the pre-images of the movement along $\mathbf{w}$ in $\mathcal{F}$.

---

4 Experiment Result

Our main purpose is to use the regularized discriminative direction to localize the class difference for human hippocampal shapes between sexes. This remains an open problem and lacks ground truth. Hence, first we have to perform a sanity check on our proposed method with data for which we know what kind of deformations to expect, and compare it with Golland's method.

The sanity check is taken on the USPS handwritten digit image database and the UMIST facial image database [5]. Each image is represented by a high-dimensional feature vector comprising all pixels, analogic to the landmark representation of shapes. The images have been known to only reside in a low-dimensional manifold [6]. We aim to discriminate (i) the shapes of two groups of digits, and (ii) two classes of human faces (8 individuals): left-side view and right-side view. In experiments, a particular feature vector is moved from one class towards the other along the discriminative direction. Note that the resulting generated images do not exist in the database. Fig. 2 is the result on USPS. As shown, Golland's method introduces much more noise (spurious difference), while our regularized method well localizes the discrimination, adding the minimum necessary shape changes. The advantages of the regularized method are more obvious on UMIST shown in Fig. 3. During deformation, it only introduces the class difference (the change of view), leaving the individual variability (the owner of the face) unchanged (Fig. 3 (c)). Most importantly, the newly generated images remain faces. However Golland's method cannot guarantee this (Fig. 3 (a)), and the authentic difference is overwhelmed by noise. Fig. 3 (b) shows the result obtained by the "tangent plane projection" (see footnote 2). It is better than Golland's method, but still worse than the regularized method (see the ghost around the glasses). This demonstrates the benefit of using $\hat{\mathbf{\Lambda}}$ in the regularized method.



(a) (b)

Fig. 2. *Sanity check on the USPS data by (a) Golland's method and (b) the regularized method. The top row shows the deformation from digit 5 to digit 8. The bottom row shows the deformation from digit 0 to digit 9.*

After the sanity check, we analyze the class difference of hippocampal shapes for sex. This is part of a longitudinal study in mental health research in Australian National University. Hand-traced left hippocampi of healthy individuals is used, which comprise 219 females and 181 males in an age span of 40-44. Each shape is normalized with respect to volume and represented by 642 landmarks generated by spherical harmonics (SPHARM) [7] with degree 5. An SVM classifier with the RBF kernel is employed for classification. The localized discrimination is shown in Fig. 4. These hippocampi belong to 6 individuals (a column for each one), 3 females and 3 males. The color code indicates the nature of deformation that an actual hippocampal shape undergoes to become a shape akin to the opposite class. Take the leftmost hippocampus in Fig. 4 for example. To make this female hippocampus to be male-like, the blue areas should shrink. As observed, the shape changes are not uniform over the whole hippocampus: small

Fig. 3. *Sanity check on UMIST data. (a) Golland's method, (b) the tangent plane projection (see footnote 3), (c) the regularized method. During the deformation, a right-side view face (the leftmost image) gradually turns towards the left (keeping adding class difference) while remaining a face image of the same person (filtering individual variability) in (b) and (c). However Golland's method in (a) cannot guarantee this.*

changes (either compression or expansion, in green color) occur on most of the shape, while sharp changes are localized on the head and the tail. Comparing the deformations in both ways (female to male and vice versa), the regularized method consistently captures the compression in the lateral parts at the head and the tail for male hippocampi. Compared with Golland's method which shows a different pattern (a compression next to an expansion in the head), our results are also more compact, with changes concentrated in fewer regions but at greater magnitude. Interestingly, the work in [4] has reported findings similar to that of our method. In [4], the hippocampal shapes are represented by medial models, totally different from our SPHARM-based shape descriptors. Shape difference for sex is observed and found that it is mostly due to the volume loss in males with age in young adulthood in the lateral areas of the hippocampus head and tail, which is not observed for females. This finding supports that of our regularized method.

5 Conclusion

Our research demonstrates the importance and benefit of incorporating the shape distribution in identifying the essential difference between two shape classes. The proposed regularized discriminative direction is applied to studying the sex difference in hippocampal shapes, localizing the key difference at the lateral parts of the head and tail. More applications are expected in our future work.

References

1. Polina Golland: Discriminative direction for kernel classifiers. In: Advances in Neural Information Processing Systems (NIPS). (2001) 745–752

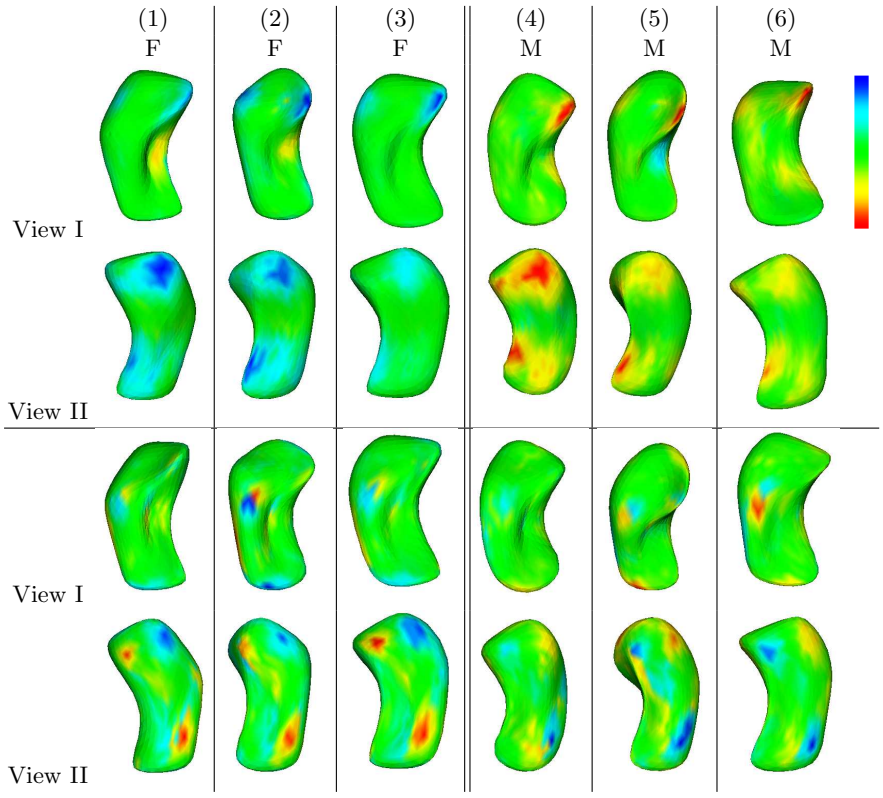


Fig. 4. *Localized discrimination for sex on hippocampi of 6 individuals (three females on the left, three males on the right) from two perspective of views. The top two rows are generated by our regularized method, while the bottom two rows are generated by Golland's method. The color code indicates the deformation that a female/male hippocampus undergoes to become a male-like/female-like one. Green indicates small shape change. From green to red, the amount of protrusion increases. From green to blue, the amount of shrinkage increases.*

2. P. Golland, W.E. Grimson, M.E. Shenton, R. Kikinis: Detection and analysis of statistical differences in anatomical shape. Med. Image Analysis **9**(1) (2005) 69–85
3. L. Zhou, R. Hartley, P. Lieby, N. Barnes, K. Anstey, N. Cherbuin, P. Sachdev: A study of hippocampal shape difference between genders by efficient hypothesis test and discriminative deformation. In: Proceedings of MICCAI. (2007) 375–383
4. Bouix S., Pruessner J. C., Collins D. L., Siddiqi K.: Hippocampal shape analysis using medial surfaces. NeuroImage **25** (2005) 1077–1089
5. D. B. Graham, N. M. Allinson: Face recognition: From theory to applications. NATO ASI Series F, Computer and Systems Sciences **163** (1998) 446–456
6. Tenenbaum J. B., de Silva V., Langford J. C.: A global geometric framework for nonlinear dimensionality reduction. Science **290** (2000) 2319–2323
7. A. Kelemen, G. Szekely, G. Gerig: Elastic model-based segmentation of 3-D neuroradiological data sets. IEEE Trans. on Medical Imaging **18**(10) (1999) 828–839