Fiducial Detection

Diane Lu

October 17, 2019

Deformable Shape Detection

Goal: To detect the positions of a dense set of specific shape landmarks, or *fiducials*, in an image.

Figure 1: Shape Detection Applications

Problems with Existing Solutions

- Require 3D database
- Find only salient landmarks

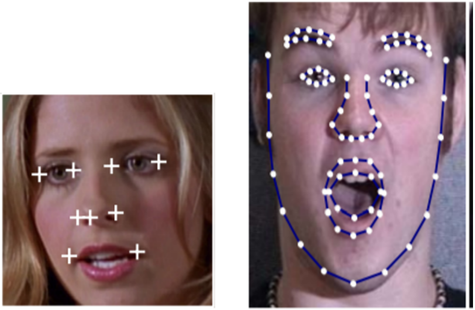

Figure 2: Only detect salient landmarks

Proposed Method

Goal: To detect a dense set of salient and non-salient fiducials in an image using 2D image data set.

Idea: Build a probabilistic graphical model for the fiducials' positions, and then estimate by MLE.

Reference: Benitez-Quiroz et al. (2014)

Workflow

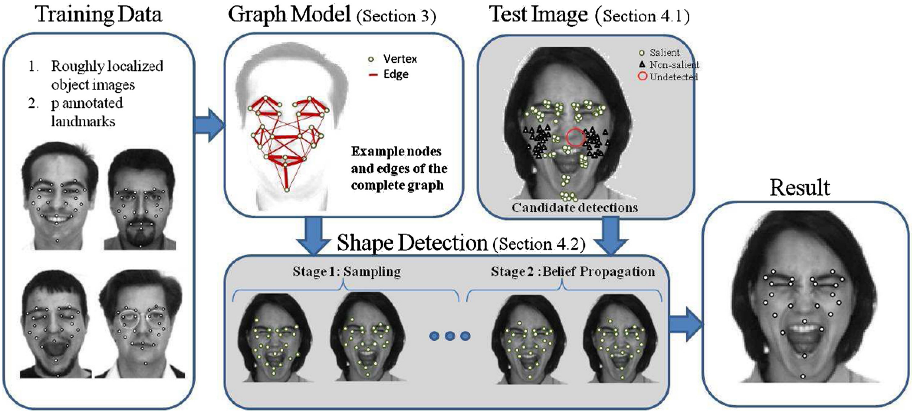

Figure 3: Workflow of the proposed method

Probabilistic Graph Model

Let x_i be 2D coordinates of fiducial i with a total of p fiducials and denote a set of coordinates as $X = (x_1, ..., x_p)$.

$$P(X) = \frac{1}{Z}\beta(X) \prod_{i=1,j=1}^{i=p,j<i} \phi_{ij}(x_i, x_j) \prod_{k=1}^{p} \gamma_k(x_k)$$

$\beta(X)$: global configuration (shape)

$\phi(.)$: relationship between two fiducials

$\gamma(.)$: texture/appearance information

Z is the normalizing constant.

Global Configuration

$$\beta(X) = e^{-\frac{\alpha}{2}(T(X)-\mu)^T \sum^{-1}(T(X)-\mu)}$$

T(X) is a normalizing transformation.

$\mu, \sum$ are the sample mean and covariance of T(X) in the training samples.

Relationship Between Two Fiducials

$$\phi_{ij}(x_i, x_j) = e^{-(1-\alpha)\overline{w}_{ij}(\Delta_{ij}-\mu_{ij})^T \sum_{ij}^{-1}(\Delta_{ij}-\mu_{ij})}$$

$\Delta_{ij} = x_i - x_j$: pair wise differences between fiducial i and j.

$\alpha \in [0, 1]$: penalty of differing from mean shape.

$\overline{w}_{ij} = \frac{1}{||\sum_{ij}||_F}$: normalized relative importance of edges connecting fiducial i and j.

Texture

γ_k : k-th fiducial's local texture = normalized confidences of local detection.

Salient landmarks can be detected reliably just through texture information. But non-salient landmarks can't.

Local Detection

Evaluating a classifier for each fiducial at each image patch within the region expected to contain the specific fiducial.

Classifier: kernel Fisher discriminant analysis (KFDA or KLDA)

Outputs a set of candidate positions D_i for each fiducial.

To be robust to occulusion, $\tilde{D}_i = \{D_i, x_i^{MLE}\}$ where $x_i^{MLE} = argmax_{x_i} \beta(X) \prod_{j=1}^{j<i} \phi_{ij}(x_i, x_j)$

Sampling and Belief Propogation

Problem: Search space for maximizing P(X) might be too large.

Solution: Sampling and belief propogation

1. Gibbs sampling: $x_i^t \sim P(x_i | x_{-i}^{t-1})$ where x_i^0 is initialized by a random sample from the $\tilde{D}_i$.
2. Belief propagation: $x_i^{t'} = argmax_{x_i \in \tilde{D}_i} P(x_i | x_{-i}^{t'-1})$ where $x^0 = argmax_{X_i^t} P(X_i^t)$

Through experiments, maxIter for t' and t is 100.

Experiments

Face Data Set	AR	LFW	XM2VTS
Training Size (images)	448	1027	448
Testing Size (images)	448	500	350

Image size: 150x150 pixels.

Number of fiducials: 50

Face Landmark Detection

Figure 4: Results on 3 data sets

Face Landmark Detection

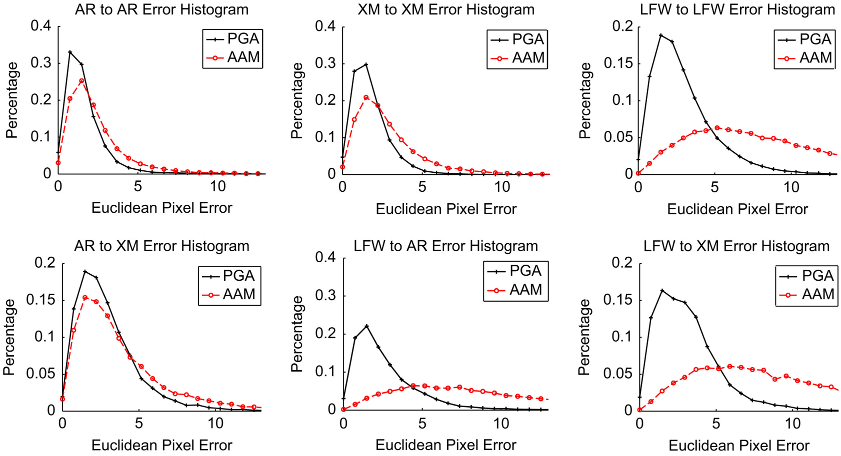

Figure 5: Error histograms. PGA is the proposed method and AAM is a method that utilizes global configuration and texture.

Cardiac MRI

Data set:
160 MRI images of left-ventricle from 8 subjects.
100 training images and 60 testing images.

Image size: 100x100 pixels

Number of landmarks: 22

Goal: Detect contour of left-venricle.

Cardiac MRI

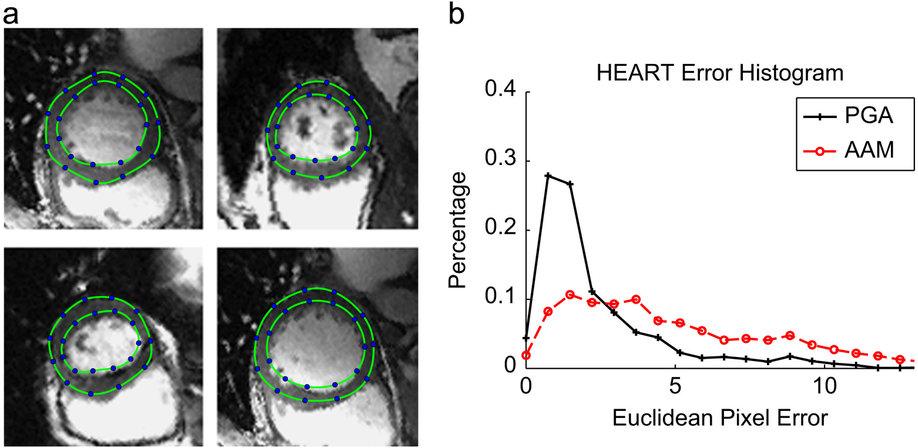

Figure 6: Heart experiment

Hand Shapes

Data set:
40 images of 4 subjects' hands.
3-fold CV, each fold has 30 training images and 10 testing images.

Image size: 100x100 pixels

Number of landmarks: 52

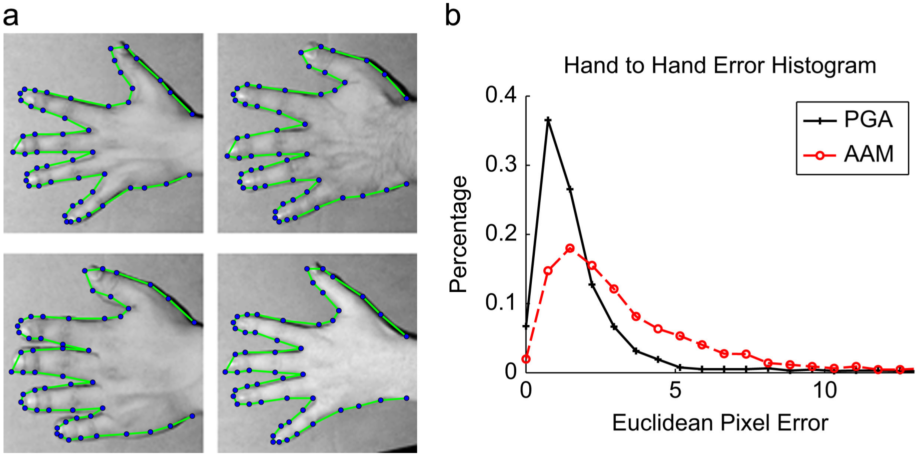

Figure 7: Hand experiment

Inferring Additional Landmarks (Incremental Learning)

Goal: Infer m fiducials from p fiducials (m $>>$ p).

Idea: Detected landmarks constrain the positions of undetected ones.

Steps:

1. Learn the parameters for m node graph.
2. $x_i^{MLE} = argmax_{x_i} \beta(X) \prod_{j=1}^{j<=p} \phi_{ij}(x_i, x_j) \; \forall i \in \{p+1, ..., m\}$

Inferring Additional Landmarks (Incremental Learning)

Figure 8: From 50 fiducials to 132 fiducials

Takeaway

Exploiting the fact that each landmark positions constrains the position of other landmarks, a probablistic graph model is built for accurately detecting a denset set of salient and non-salient fiducials.

The face experiment shows that the algorithm can detect significant deformations of eyes and mouth.

Also, the proposed method can be applied to many other fields (hand and heart experiments).

References I

Benitez-Quiroz, Carlos F, Samuel Rivera, Paulo FU Gotardo, and Aleix M Martinez. 2014. "Salient and Non-Salient Fiducial Detection Using a Probabilistic Graphical Model." *Pattern Recognition* 47 (1). Elsevier: 208–15.