Expectation Maximization Algorithm

by José Ananías Hilario Reyes

I. Mathematical Analysis

We have the following basic definitions

$\mathcal{T}^m = \{\mathbf{x}^1, \ldots, \mathbf{x}^m\}$

$D \subset \mathbb{Z}^2$

$i \in D$

$\mathbf{s} : D \mapsto \{0,1\}$ where s_i denotes the segment label in pixel $i \in D$

$\mathbf{x} : D \mapsto \mathbb{R}^3$ where x_i denotes the three dimensional colour vector in pixel $i \in D$

$\theta^0, \theta^1, \ell = 1, 2, \ldots, m$

$$p_{\mathbf{u}}(\mathbf{s}) = \prod_{i \in D} p_{u_i}(s_i) = \prod_{i \in D} \frac{e^{u_i S_i}}{1 + e^{u_i}}$$

$$p_{u_i}(s_i = 0) = \frac{1}{1 + e^{u_i}}$$

$$p_{u_i}(s_i = 1) = \frac{e^{u_i}}{1 + e^{u_i}}$$

$$p_{\boldsymbol{\theta}}(x_i | s_i = \{0,1\}) \sim \mathcal{N}(\mu, \Sigma)$$

$$\ell(\mathbf{u}, \boldsymbol{\theta}) = \prod_{l=1}^{m} \prod_{i \in D} p_{\mathbf{u},\boldsymbol{\theta}}(x_i, s_i)$$

$$\ell(\mathbf{u}, \boldsymbol{\theta}) = \prod_{l=1}^{m} \prod_{i \in D} p_{\mathbf{u}}(s_i) p_{\theta^\ell}(x_i^\ell | s_i)$$

$$p_{\mathbf{u},\boldsymbol{\theta}}(x_i, s_i) = p_{\mathbf{u}}(s_i) p_{\boldsymbol{\theta}}(x_i|s_i) = \sum_{s_i \in \{0,1\}} p_{u_i}(s_i) p_{\boldsymbol{\theta}}(x_i|s_i)$$

$$\sum_{s_i \in \{0,1\}} p_{u_i}(s_i) p_{\boldsymbol{\theta}}(x_i|s_i) = p_{u_i}(s_i = 0) p_{\boldsymbol{\theta}}(x_i|s_i = 0) + p_{u_i}(s_i = 1) p_{\boldsymbol{\theta}}(x_i|s_i = 1)$$

$$\sum_{s_i \in \{0,1\}} p_{u_i}(s_i) p_{\boldsymbol{\theta}}(x_i|s_i) = \frac{p_{\boldsymbol{\theta}}(x_i|s_i = 0)}{1 + e^{u_i}} + \frac{e^{u_i} p_{\boldsymbol{\theta}}(x_i|s_i = 1)}{1 + e^{u_i}}$$

Then

$$p_{\mathbf{u},\boldsymbol{\theta}}(x_i, s_i) = \sum_{s_i \in \{0,1\}} p_{u_i}(s_i) p_{\boldsymbol{\theta}}(x_i|s_i) = \sum_{s_i \in \{0,1\}} p_{u_i,\boldsymbol{\theta}}(x_i, s_i)$$

$$\ell(\mathbf{u}, \boldsymbol{\theta}) = \prod_{l=1}^{m} \prod_{i \in D} p_{\mathbf{u},\boldsymbol{\theta}}(x_i, s_i) = \prod_{l=1}^{m} \prod_{i \in D} \sum_{s_i \in \{0,1\}} p_{u_i,\boldsymbol{\theta}}(x_i, s_i)$$

$$\boxed{\mathcal{L}(\mathbf{u}, \boldsymbol{\theta}) = \log \prod_{l=1}^{m} \prod_{i \in D} p_{\mathbf{u},\boldsymbol{\theta}}(x_i, s_i) = \sum_{l=1}^{m} \sum_{i \in D} \log p_{\mathbf{u},\theta^\ell}(x_i^\ell, s_i) = \sum_{l=1}^{m} \sum_{i \in D} \log \sum_{s_i \in \{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i)}$$

Now, by Jensen’s inequality

$$\log \sum_{s_i \in \{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i) \geq \sum_{s_i \in \{0,1\}} \alpha_\ell(s_i) \log p_{u_i,\theta^\ell}(x_i^\ell, s_i) - \sum_{s_i \in \{0,1\}} \alpha_\ell(s_i) \log \alpha_\ell(s_i)$$

We get the objective function for the EM-Algorithm

$$\frac{1}{m} \sum_{\ell=1}^{m} \sum_{i \in D} \sum_{s_i \in \{0,1\}} \left[\alpha_\ell(s_i) \log p_{u_i,\theta^\ell}(x_i^\ell, s_i) - \alpha_\ell(s_i) \log \alpha_\ell(s_i)\right] \to \underset{\mathbf{u},\boldsymbol{\theta},\alpha}{\arg\max}$$

Deduce that the E-step reads as simple as

$$\alpha_\ell(s_i) = p_{u_i,\theta^\ell}(s_i|x_i^\ell)$$

Recalling log-likelihood function

$$\mathcal{L}(\mathbf{u}, \boldsymbol{\theta}) = \sum_{l=1}^{m} \sum_{i \in D} \log \sum_{s_i \in \{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i)$$

We can add the following fraction as it doesn't change the value of the equation

$$\sum_{l=1}^{m} \sum_{i \in D} \log \sum_{s_i \in \{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i) = \sum_{l=1}^{m} \sum_{i \in D} \log \sum_{s_i \in \{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i) \frac{\alpha_\ell(s_i)}{\alpha_\ell(s_i)}$$

Then, we can see the inner sum as an expectation

$$\log \sum_{s_i \in \{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i) \frac{\alpha_\ell(s_i)}{\alpha_\ell(s_i)} = \log \sum_{s_i \in \{0,1\}} \alpha_\ell(s_i) \frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)} = \log \mathbb{E}\left[\frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right]$$

By Jensen's inequality we know that

$$\log \mathbb{E}\left[\frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right] \geq \mathbb{E}\left[\frac{\log p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right]$$

From where we can then move the multiplying $\alpha_\ell(s_i)$ out of the $\log$

$$\mathbb{E}\left[\frac{\log p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right] = \alpha_\ell(s_i) \sum_{s_i \in \{0,1\}} \log \frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}$$

Summarizing

$$\log \sum_{s_i \in \{0,1\}} \frac{\alpha_\ell(s_i) p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)} \geq \alpha_\ell(s_i) \sum_{s_i \in \{0,1\}} \log \frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}$$

Now we have a lower bound of objective function which is tractable. Furthermore, we want the inequality to hold with equality, so that we can solve the optimization problem by coordinate ascent.

We will do this by choosing $\alpha_\ell(s_i)$ such that

$$\log \mathbb{E}\left[\frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right] = \mathbb{E}\left[\frac{\log p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right]$$

Which means that

$$\frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)} = \mathbb{E}\left[\frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)}\right]$$

$$\frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\alpha_\ell(s_i)} \rightarrow \text{ is a constant!}$$

Then

$$\alpha_\ell(s_i) \propto p_{u_i,\theta^\ell}(x_i^\ell, s_i)$$

$$\sum_{s_i=\{0,1\}} \alpha_\ell(s_i) = 1$$

$$\boxed{\alpha_\ell(s_i) = \frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{\sum_{s_i=\{0,1\}} p_{u_i,\theta^\ell}(x_i^\ell, s_i)} = \frac{p_{u_i,\theta^\ell}(x_i^\ell, s_i)}{p_{u_i,\theta^\ell}(x_i^\ell)} = p_{u_i,\theta^\ell}(s_i|x_i^\ell)}$$

$$\boxed{\alpha_\ell(s_i) = p_{u_i,\theta^\ell}(s_i|x_i^\ell)}$$

Deduce that the <u>M-step</u> decomposes into two independent optimization tasks

a) **Show that the maximization wrt $\mathbf{u}$ further decomposes into independent tasks for each pixel $i \in D$:**

$$\frac{1}{m}\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)u_i - \log(1 + e^{u_i}) \rightarrow \arg\max_{u_i}$$

Show also that the function is concave and has a unique global minimum.

$$\frac{\partial}{\partial u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\sum_{i\in D}\sum_{s_i\in\{0,1\}} [\alpha_\ell(s_i)\log p_{u_i,\theta^\ell}(x_i^\ell, s_i) - \alpha_\ell(s_i)\log\alpha_\ell(s_i)]\right] = 0$$

$$\frac{\partial}{\partial u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\sum_{i\in D}\sum_{s_i\in\{0,1\}} \alpha_\ell(s_i)\left[\log p_{\theta^\ell}(x_i^\ell|s_i)p_{u_i}(s_i) - \log\alpha_\ell(s_i)\right]\right] = 0$$

Because $p_{\mathbf{u}}(\mathbf{s})$ is the same for all images, but varies between pixels, the optimization task is carried out per pixel. Therefore, we can drop $\sum_{i \in D}$

$$\frac{\partial}{\partial u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\sum_{s_i\in\{0,1\}}\alpha_\ell(s_i)\left[\cancelto{0}{\log p_{\theta^\ell}(x_i^\ell|s_i)} + \log p_{u_i}(s_i) - \cancelto{0}{\log \alpha_\ell(s_i)}\right]\right]$$

$$\frac{\partial}{\partial u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\sum_{s_i\in\{0,1\}}\alpha_\ell(s_i)\log p_{u_i}(s_i)\right] = 0$$

For $s_i = 0$:

$$\alpha_\ell(s_i=0)\log\left(\frac{1}{1+e^{u_i}}\right) = \alpha_\ell(s_i=0)e - \alpha_\ell(s_i=0)\log(1+e^{u_i})$$

$$\frac{\partial}{\partial u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\left[\cancelto{0}{\alpha_\ell(s_i=0)e} - \alpha_\ell(s_i=0)\log(1+e^{u_i})\right]\right] = 0$$

For $s_i = 1$:

$$\alpha_\ell(s_i=1)\log\left(\frac{e^{u_i}}{1+e^{u_i}}\right) = \alpha_\ell(s_i=1)u_i - \alpha_\ell(s_i=1)\log(1+e^{u_i})$$

$$\frac{\partial}{\partial u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\left[\alpha_\ell(s_i=1)u_i - \alpha_\ell(s_i=1)\log(1+e^{u_i})\right]\right] = 0$$

Merging the equations:

$$\arg\max_{u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\left[\alpha_\ell(s_i=1)u_i - \cancelto{1}{(\alpha_\ell(s_i=0)+\alpha_\ell(s_i=1))}\log(1+e^{u_i})\right]\right]$$

$$\arg\max_{u_i}\left[\frac{1}{m}\sum_{\ell=1}^{m}\left[\alpha_\ell(s_i=1)u_i - \log(1+e^{u_i})\right]\right]$$

$$\arg\max_{u_i} \left[\frac{1}{m} \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) u_i - \cancelto{1}{\frac{1}{m} \sum_{\ell=1}^{m}} \log(1 + e^{u_i}) \right]$$

$$\boxed{\arg\max_{u_i} \left[\frac{1}{m} \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) u_i - \log(1 + e^{u_i}) \right]}$$

$$\frac{1}{m} \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) - \frac{e^{u_i}}{1 + e^{u_i}} \cdot \frac{e^{-u_i}}{e^{-u_i}} = 0$$

$$\frac{1}{m} \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) - \frac{1}{1 + e^{-u_i}} = 0$$

$$\frac{1}{m} \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) = \frac{1}{1 + e^{-u_i}}$$

$$\frac{1 + e^{-u_i}}{1} = \frac{m}{\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}$$

$$1 + e^{-u_i} = \frac{m}{\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}$$

$$e^{-u_i} = \frac{m}{\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)} - 1 = \frac{m - \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}{\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}$$

$$e^{u_i} = \frac{\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}{m - \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}$$

$$\boxed{u_i = \log\left[\frac{\sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)}{m - \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1)} \right]}$$

$$\frac{\partial^2}{\partial^2 u_i} \left[\frac{1}{m} \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) - \frac{e^{u_i}}{1 + e^{u_i}} \right]$$

$$-\frac{e^{u_i}}{\left(1+e^{u_i}\right)^2} \leq 0 \Rightarrow \textbf{is Concave function}$$

As long as $0 < \sum_{\ell=1}^{m} \alpha_\ell(s_i = 1) < m$

$$\boxed{u_i = \log\left[\frac{\sum_{\ell=1}^{m} \alpha_\ell(s_i=1)}{m - \sum_{\ell=1}^{m} \alpha_\ell(s_i=1)}\right]} \Rightarrow \textbf{is global maxima}$$

b) Show that the maximization task wrt the parameters θ decomposes into independent tasks for each image, and the foreground resp. Background appearance parameters:

$$\sum_{i \in D} \alpha_\ell(s_i = 1) \log p_{\theta_1^\ell}(x_i^\ell | s_i = 1) \to \arg\max_{\theta_1^\ell}$$

$$\sum_{i \in D} \alpha_\ell(s_i = 0) \log p_{\theta_0^\ell}(x_i^\ell | s_i = 0) \to \arg\max_{\theta_0^\ell}$$

$$\frac{\partial}{\partial \theta^\ell}\left[\frac{1}{m}\sum_{\ell=1}^{m}\sum_{i \in D}\sum_{s_i \in \{0,1\}}\left[\alpha_\ell(s_i)\log p_{u_i,\theta^\ell}(x_i^\ell, s_i) - \alpha_\ell(s_i)\log\alpha_\ell(s_i)\right]\right] = 0$$

$$\frac{\partial}{\partial \theta^\ell}\left[\frac{1}{m}\sum_{\ell=1}^{m}\sum_{i \in D}\sum_{s_i \in \{0,1\}}\alpha_\ell(s_i)\left[\log p_{\theta^\ell}(x_i^\ell | s_i) p_{u_i}(s_i) - \log\alpha_\ell(s_i)\right]\right] = 0$$

Because $p_{\boldsymbol{\theta}}(x_i|s_i)$ is the same for each pixel, but varies across images, the optimization task is carried out per image. Therefore, we can drop $\frac{1}{m}\sum_{\ell=1}^{m}$

$$\frac{\partial}{\partial \theta^\ell}\left[\sum_{i \in D}\sum_{s_i \in \{0,1\}}\alpha_\ell(s_i)\left[\log p_{\theta^\ell}(x_i^\ell | s_i) + \cancelto{0}{\log p_{u_i}(s_i)} - \cancelto{0}{\log\alpha_\ell(s_i)}\right]\right] = 0$$

$$\frac{\partial}{\partial \theta^\ell}\left[\sum_{i\in D}\sum_{s_i\in\{0,1\}}\alpha_\ell(s_i)\log p_{\theta^\ell}(x_i^\ell|s_i)\right]=0$$

For $s_i = 0$:

$$\boxed{\arg\max_{\theta_0^\ell}\left[\sum_{i\in D}\alpha_\ell(s_i=0)\log p_{\theta_0^\ell}(x_i^\ell|s_i=0)\right]}$$

For $s_i = 1$:

$$\boxed{\arg\max_{\theta_1^\ell}\left[\sum_{i\in D}\alpha_\ell(s_i=1)\log p_{\theta_1^\ell}(x_i^\ell|s_i=1)\right]}$$

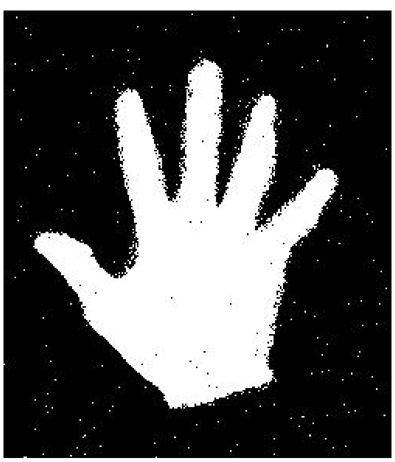
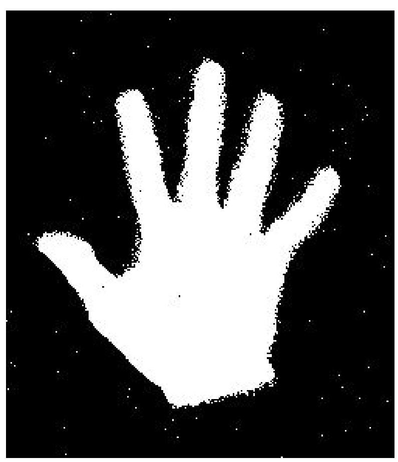
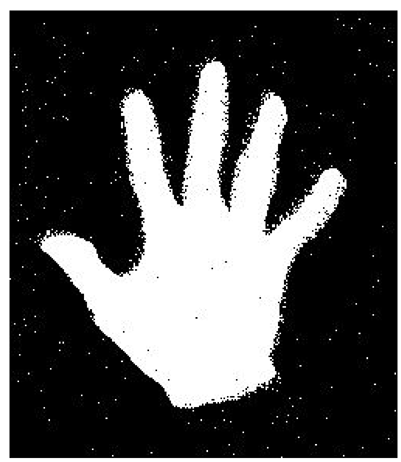
II. Baseline approaches

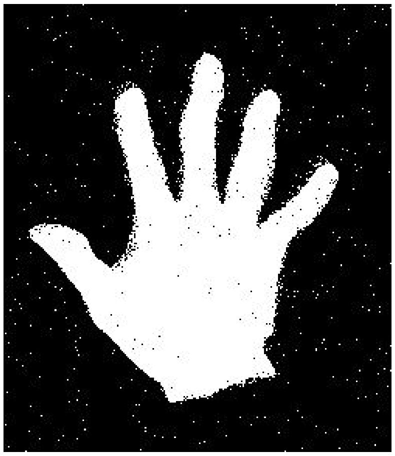
The mixture approach yields slightly better results. Gaussian Mixture model is a generalization over K-means, so it comes out as no surprise that their performance is nearly identical. However, as it is indeed a generalization over K-means, it could be expected to perform slightly better in tasks that are not entirely suited to K-means clustering by itself.

For comparison of results, refer to table in next page.

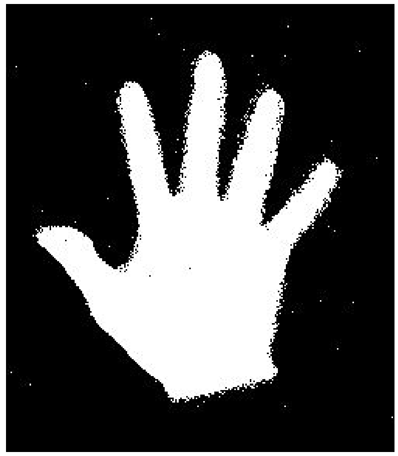
III. EM-Algorithm output model and classification

Final shape model

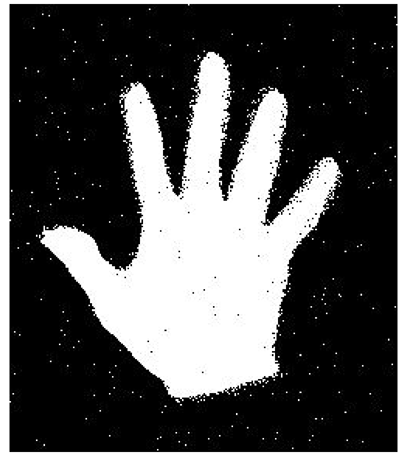
<table>
<tr><th colspan="3">Classification Examples</th></tr>
<tr><td>
hand_00</td><td>
hand_03</td><td>
hand_05</td></tr>
</table>



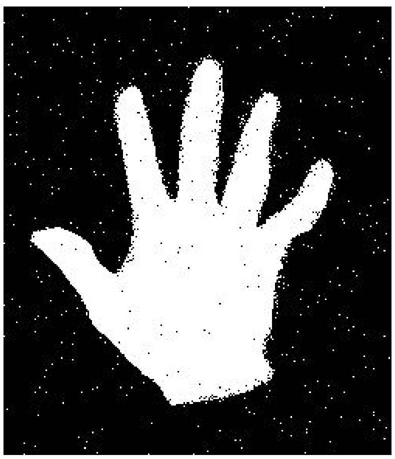
hand_09



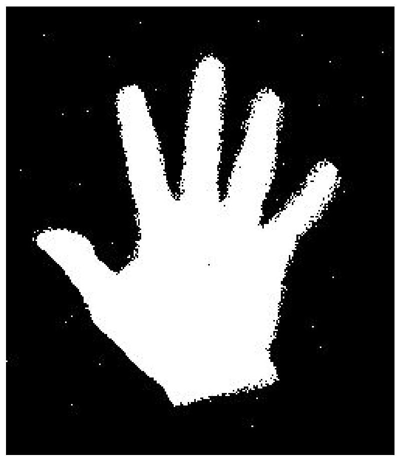
hand_16



hand_17

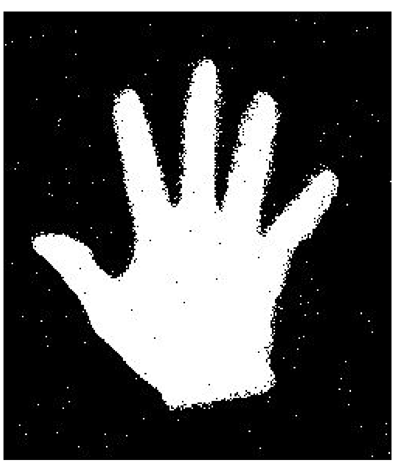


hand_22

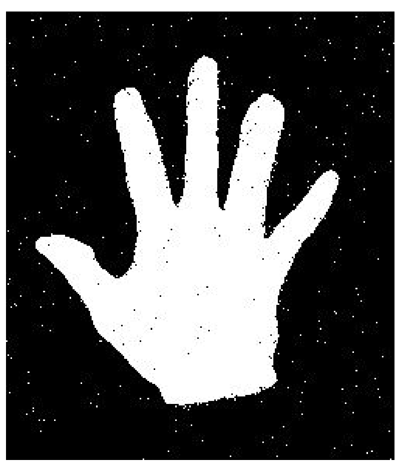


hand_27

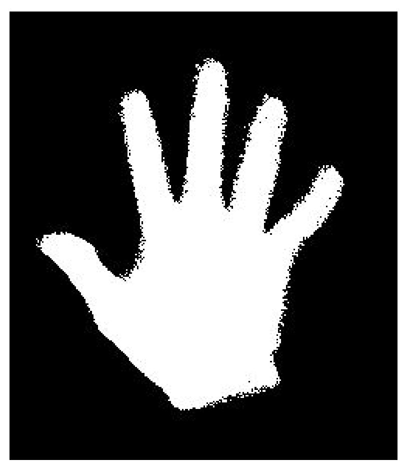
hand_30



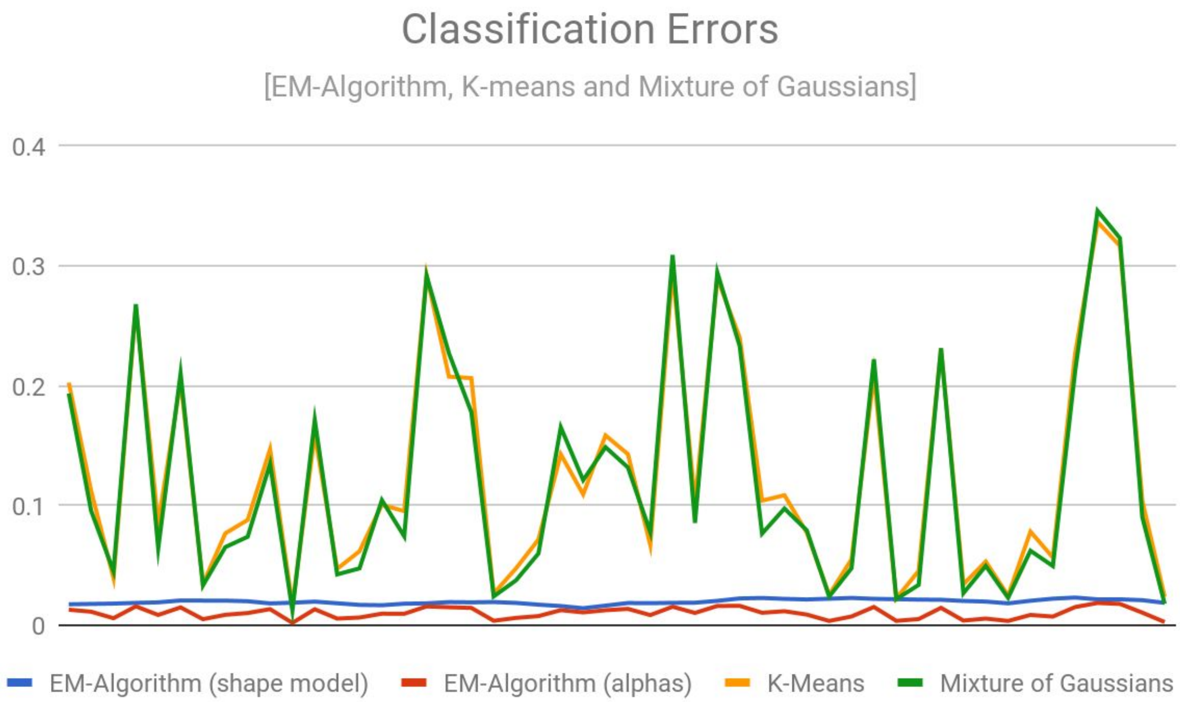
hand_39



hand_41



hand_46



EM-Algorithm (shape model)	EM-Algorithm (alphas)	K-Means	Mixture of Gaussians
0.01775778547	0.01323183391	0.2026989619	0.1936470588
0.01806228374	0.01157093426	0.1124290657	0.09580622837
0.01833910035	0.00614532872	0.0398200692	0.04537024221
0.01892041522	0.01601384083	0.2656747405	0.2682076125
0.01937716263	0.008858131488	0.08502422145	0.06451211073
0.02092733564	0.01516955017	0.2041107266	0.2102698962
0.02084429066	0.005314878893	0.03476816609	0.03353633218
0.020816609	0.008982698962	0.0769550173	0.06546712803
0.02029065744	0.01044982699	0.08812456747	0.07410380623
0.01850519031	0.01370242215	0.1470449827	0.1349342561
0.01905882353	0.002256055363	0.01674740484	0.01269204152
0.01997231834	0.01360553633	0.1598477509	0.1714878893
0.01861591696	0.00584083045	0.04734948097	0.04269896194
0.01725951557	0.006698961938	0.06196539792	0.04772318339
0.01698269896	0.009910034602	0.1007612457	0.1046366782
0.01824221453	0.00984083045	0.09548788927	0.07446366782

0.0185467128	0.01598615917	0.2924290657	0.2918615917
0.01958477509	0.01529411765	0.2076816609	0.2271280277
0.01937716263	0.01475432526	0.206449827	0.1780761246
0.01964013841	0.00401384083	0.02674048443	0.02433217993
0.01889273356	0.006394463668	0.04779238754	0.03778546713
0.0174532872	0.007903114187	0.07176470588	0.06011072664
0.01629065744	0.01284429066	0.1428927336	0.1654256055
0.01446366782	0.01098961938	0.1094948097	0.1213564014
0.0166366782	0.01280276817	0.1586574394	0.1488581315
0.01882352941	0.01392387543	0.1430865052	0.1319584775
0.01868512111	0.008761245675	0.06732179931	0.07795155709
0.01897577855	0.01572318339	0.2958062284	0.3093148789
0.01908650519	0.0105467128	0.1026989619	0.08564705882
0.02060899654	0.01637370242	0.2897024221	0.2945882353
0.02262975779	0.01644290657	0.2403044983	0.2331211073
0.02301730104	0.01068512111	0.1042491349	0.07676124567
0.02226989619	0.01197231834	0.1087612457	0.09764705882
0.02178546713	0.009162629758	0.07749480969	0.07943252595
0.02239446367	0.003875432526	0.02568858131	0.02420761246
0.0230449827	0.007584775087	0.0553633218	0.04790311419
0.02229757785	0.01551557093	0.2096055363	0.222200692
0.02199307958	0.003944636678	0.02262975779	0.02206228374
0.02167474048	0.005425605536	0.0455916955	0.03393771626
0.02152249135	0.01480968858	0.2313217993	0.2314463668
0.02055363322	0.004179930796	0.03443598616	0.02703114187
0.02002768166	0.005826989619	0.05341176471	0.04975778547
0.0185467128	0.003861591696	0.02489965398	0.02319723183
0.02070588235	0.008871972318	0.0784083045	0.06243598616
0.02243598616	0.007501730104	0.056816609	0.04970242215
0.0233633218	0.01546020761	0.227349481	0.2125951557
0.02186851211	0.01885121107	0.336733564	0.3460761246
0.0219100346	0.0180899654	0.3168027682	0.3232802768
0.02117647059	0.01080968858	0.1040276817	0.09050519031
0.01897577855	0.002948096886	0.02395847751	0.01814532872