Let F denote the female training set and M the male training set. Let $d = |F| + |M|$, and let N^2 be the number of pixels in each image. Determining the optimal decision boundary requires finding $\mathbf{w} \in \mathbb{R}^{N^2}$ satisfying $\mathbf{Sw} = \mathbf{m}_F - \mathbf{m}_M$, where $\mathbf{m}_F$, $\mathbf{m}_M \in \mathbb{R}^{N^2}$ are the means of F and M, respectively, and $\mathbf{S} \in \mathbb{R}^{N^2 \times N^2}$ is the within-class scatter matrix given by

$$\mathbf{S} = \sum_{\mathbf{x} \in F} (\mathbf{x} - \mathbf{m}_F)(\mathbf{x} - \mathbf{m}_F)^{\mathrm{T}} + \sum_{\mathbf{x} \in M} (\mathbf{x} - \mathbf{m}_M)(\mathbf{x} - \mathbf{m}_M)^{\mathrm{T}}.$$

(For brevity we use the notation $\mathbf{S}$ instead of $\mathbf{S}_W$.) In our case, $N = 256$, so representing $\mathbf{S}$ in a computer is cumbersome (indeed, disallowed by MATLAB's standard settings). We seek a means to solve $\mathbf{Sw} = \mathbf{m}_F - \mathbf{m}_M$ without computing $\mathbf{S}$ explicitly.

We may represent $\mathbf{S}$ by $\mathbf{S} = \mathbf{CC}^{\mathrm{T}}$, where $\mathbf{C} \in \mathbb{R}^{N^2 \times d}$ is the matrix whose i^{th} column is the i^{th} element of $F \cup M$. $\mathbf{C}$ has the singular value decomposition $\mathbf{C} = \mathbf{U\Sigma V}^{\mathrm{T}}$, where

- $\mathbf{U} \in \mathbb{R}^{N^2 \times N^2}$ is a unitary matrix whose i^{th} column is the i^{th} (normalized) eigenvector of $\mathbf{S}$;
- $\mathbf{\Sigma} \in \mathbb{R}^{N^2 \times d}$ satisfies $\sigma_{i,i} = \sqrt{\lambda_i}$ for $1 \leq i \leq d$, where λ_i is the i^{th} (necessarily non-negative) eigenvalue of $\mathbf{C}^{\mathrm{T}}\mathbf{C} \in \mathbb{R}^{d \times d}$, and $\sigma_{i,j} = 0$ when $i \neq j$;
- $\mathbf{V} \in \mathbb{R}^{d \times d}$ is a unitary matrix whose i^{th} column is the i^{th} (normalized) eigenvector of $\mathbf{C}^{\mathrm{T}}\mathbf{C} \in \mathbb{R}^{d \times d}$.

Henceforth let $\mathbf{B} = \mathbf{C}^{\mathrm{T}}\mathbf{C}$.

Therefore,

$$\mathbf{S} = (\mathbf{U\Sigma V}^{\mathrm{T}})(\mathbf{U\Sigma V}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{U\Sigma V}^{\mathrm{T}}\mathbf{V}(\mathbf{U\Sigma})^{\mathrm{T}} = \mathbf{U\Sigma}(\mathbf{U\Sigma})^{\mathrm{T}},$$

and thus we seek $\mathbf{w}$ satisfying

$$\begin{aligned}
\mathbf{U\Sigma}(\mathbf{U\Sigma})^{\mathrm{T}}\mathbf{w} &= \mathbf{m}_F - \mathbf{m}_M \\
\mathbf{U}^{\mathrm{T}}\mathbf{U\Sigma}(\mathbf{U\Sigma})^{\mathrm{T}}\mathbf{w} &= \mathbf{U}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
\mathbf{\Sigma}(\mathbf{U\Sigma})^{\mathrm{T}}\mathbf{w} &= \mathbf{U}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}(\mathbf{U\Sigma})^{\mathrm{T}}\mathbf{w} &= \mathbf{\Sigma}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
(\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma})(\mathbf{U\Sigma})^{\mathrm{T}}\mathbf{w} &= (\mathbf{U\Sigma})^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M).
\end{aligned}$$

For i from 1 to d, the $(i,i)^{\text{th}}$ entry of $\mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma} \in \mathbb{R}^{d \times d}$ is $\sigma_i^2 = \lambda_i$; all other entries are 0. In this spirit, let $\mathbf{\Lambda} = \mathbf{\Sigma}^{\mathrm{T}}\mathbf{\Sigma}$. Under the reasonable assumption that no eigenvalue of $\mathbf{B}$ is 0, $\mathbf{\Lambda}$ is invertible, and

$$(\mathbf{U\Sigma})^{\mathrm{T}}\mathbf{w} = \mathbf{\Lambda}^{-1}(\mathbf{U\Sigma})^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M).$$

Let $\mathbf{A} = \mathbf{U\Sigma} \in \mathbb{R}^{N^2 \times d}$. For i from 1 to d, let $\mathbf{e}_i$ be the i^{th} eigenvalue of $\mathbf{B}$. Then

$$\begin{aligned}
\mathbf{C}^{\mathrm{T}}\mathbf{C}\mathbf{e}_i &= \lambda_i \mathbf{e}_i \\
\mathbf{CC}^{\mathrm{T}}\mathbf{C}\mathbf{e}_i &= \lambda_i \mathbf{C}\mathbf{e}_i \\
\mathbf{S}(\mathbf{C}\mathbf{e}_i) &= \lambda_i (\mathbf{C}\mathbf{e}_i).
\end{aligned}$$

Thus $\mathbf{C}\mathbf{e}_i$ is the i^{th} eigenvector of $\mathbf{S}$, and is hence the i^{th} column of $\mathbf{U}$. (For $i > d$, the i^{th} column of $\mathbf{U}$ is a basis vector of the kernel of $\mathbf{S}$.) Therefore, the i^{th} column of $\mathbf{A}$ is $\sigma_i \mathbf{C}\mathbf{e}_i = \sqrt{\lambda_i}\mathbf{C}\mathbf{e}_i$.

Without loss of generality we may assume that $\mathbf{w} \perp \ker \mathbf{S}$, because we are trying to solve $\mathbf{S}\mathbf{w} = \mathbf{m}_F - \mathbf{m}_M$. Therefore, there exists $\mathbf{a} \in \mathbb{R}^d$ such that $\mathbf{w} = \sum_{i=1}^d a_i \mathbf{C}\mathbf{e}_i$. Therefore,

$$\begin{aligned}
\mathbf{A}^{\mathrm{T}} \sum_{i=1}^{d} a_i \mathbf{C}\mathbf{e}_i &= \boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
\mathbf{A}^{\mathrm{T}}\mathbf{C} \sum_{i=1}^{d} a_i \mathbf{e}_i &= \boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
\sum_{i=1}^{d} a_i \mathbf{e}_i &= (\mathbf{A}^{\mathrm{T}}\mathbf{C})^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M);
\end{aligned}$$

this operation is valid because

$$\mathbf{A}^{\mathrm{T}}\mathbf{C} = (\mathbf{U}\boldsymbol{\Sigma})^{\mathrm{T}}(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}) = \boldsymbol{\Sigma}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}} = \boldsymbol{\Sigma}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}} = \boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}},$$

which as the product of invertible matrices is invertible. Thus,

$$\begin{aligned}
\sum_{i=1}^{d} a_i \mathbf{e}_i &= (\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}})^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
\mathbf{C}\sum_{i=1}^{d} a_i \mathbf{e}_i &= \mathbf{C}(\boldsymbol{\Lambda}^2\mathbf{V}^{\mathrm{T}})^{-1}\mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \\
\mathbf{w} &= \mathbf{C}(\boldsymbol{\Lambda}^2\mathbf{V}^{\mathrm{T}})^{-1}\mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M).
\end{aligned}$$

We therefore have the following algorithm for finding $\mathbf{w}$:

1. Define $\mathbf{C} \in \mathbb{R}^{N^2 \times d}$: For $1 \le i \le d$, the i^{th} column is the i^{th} element of $F \cup M$.
2. Compute $\mathbf{B} = \mathbf{C}^{\mathrm{T}}\mathbf{C} \in \mathbb{R}^{d \times d}$.
3. Find the eigenvalues and eigenvectors of $\mathbf{B}$. Store the eigenvalues as the diagonal entries in $\boldsymbol{\Lambda} \in \mathbb{R}^{d \times d}$ and the (normalized) eigenvectors as the columns of $\mathbf{V} \in \mathbb{R}^{d \times d}$.
4. Define $\mathbf{A} \in \mathbb{R}^{N^2 \times d}$: For $1 \le i \le d$, the i^{th} column is
$$\frac{\sqrt{\lambda_i}}{\|\mathbf{C}\mathbf{e}_i\|_2}\mathbf{C}\mathbf{e}_i,$$
where $\mathbf{e}_i$ is the i^{th} column of $\mathbf{V}$ and λ_i is the $(i,i)^{\text{th}}$ entry of $\boldsymbol{\Lambda}$.
5. Compute $\mathbf{y} = \mathbf{A}^{\mathrm{T}}(\mathbf{m}_F - \mathbf{m}_M) \in \mathbb{R}^d$.
6. Solve $(\boldsymbol{\Lambda}^2\mathbf{V}^{\mathrm{T}})\mathbf{z} = \mathbf{y}$ for $\mathbf{z} \in \mathbb{R}^d$. The matrix $\boldsymbol{\Lambda}^2\mathbf{V}^{\mathrm{T}}$ is $d \times d$ and nonsingular, so this should be computationally inexpensive.
7. Compute $\mathbf{w} = \mathbf{C}\mathbf{z}$. Rescale if desired.