1 Basis

Een digitale foto bestaat uit *pixels*, dit zijn kleine puntjes die één bepaalde kleur hebben. Bij aan zwart-wit foto is die kleur een grijswaarde, die op de computer voorgesteld wordt door een natuurlijk getal tussen 0 en 255.

In het secundair heb je geleerd over vectoren in het vlak en in de 3-dimensionale ruimte. Je hebt geleerd dat je de **afstand** tussen twee vectoren $\mathbf{x}$ en $\mathbf{y}$ kan berekenen als

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \text{ of } d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \,.$$

We kunnen dit principe uitbreiden naar digitale foto's, die bestaan uit n pixels. Dan kunnen we zeggen dat die foto overeenkomt met een "vector" met n componenten. De afstand tussen twee foto's kunnen we dan berekenen als

$$d = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \,.$$

Stel nu dat we een databank hebben met daarin een aantal foto's $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$ en dat we een nieuwe foto $\mathbf{a}$ krijgen waarvan we moeten uitzoeken op welke foto in onze database hij het meeste lijkt. Een eenvoudige methode om dit te doen is voor elke foto in onze databank de afstand te berekenen tussen die foto in de databank en de nieuwe foto $\mathbf{a}$. De foto waarvoor die afstand het kleinste is, is dan met de grootste kans de juiste foto.

Een meer gesofisticeerde methode gaat als volgt. We weten dat de pixels van een foto van een gezicht niet zomaar willekeurige waarden aannemen. Er zijn bepaalde verbanden tussen pixels onderling. Bijvoorbeeld: alle pixels die net boven de ogen staan, vormen samen de wenkbrauwen; ze hebben typisch allemaal bijna dezelfde, vaak donkere, kleur. De pixels die samen de neus vormen, hebben meestal 2 symmetrische donkere vlekken voor de neusgaten. Enz. Om deze eigenschappen in rekening te brengen, maakt men gebruik van *eigengezichten*.

2 Eerste behandeling

We haalden al eerder aan dat elke foto $\mathbf{a}$ kan gezien worden als een vector met n componenten. Een foto is m.a.w. een vector in een n-dimensionale vectorruimte. In een vectorruimte kan je een *basis* definiëren, bestaande uit n basisvectoren $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Indien deze n vectoren aan bepaalde voorwaarden voldoen, dan bestaan er voor elke $\mathbf{a}$ getallen $c_1, c_2, \ldots, c_n$ zodat:

$$\mathbf{a} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$$

Deze getallen noemen we de *coördinaten* van $\mathbf{a}$.

n getallen vinden we teveel rekenwerk. Daarom beperken we ons tot slechts K vectoren en schrijven:

$$\mathbf{a} \approx c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_K\mathbf{u}_K$$

Het vergelijken van twee foto's kan nu gebeuren door de **afstand** tussen slechts K coördinaten van elke foto te berekenen i.p.v. tussen de n pixels. Een foto $\mathbf{a}$ is nu ook slechts benaderend voorgesteld. Het is dan ook belangrijk om goede basisvectoren te kiezen die de verschillen tussen de foto's zo goed mogelijk beschrijven.

3 Verfijning

Hier leggen we uit hoe we de geschikte vectoren $\mathbf{u}_1, \ldots, \mathbf{u}_K$ (de eigengezichten) kunnen vinden. Om enkel de relevante veschillen tussen de foto's te vatten halen we eerst het gemiddelde van alle foto's weg. We berekenen eerst de gemiddelde foto

$$\mathbf{g} = \frac{1}{M}(\mathbf{x} + \mathbf{y} + \ldots)$$

waarbij M het aantal foto's is. Daarna trekken we van elke foto dit gemiddelde af: $\mathbf{x}' = \mathbf{x} - \mathbf{g}, \ldots$. We definiëren dan de matrix $\mathbf{A}$ als de matrix met als kolommen de vectoren (de foto's min de gemiddelde foto):

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}' & \mathbf{y}' & \ldots \end{bmatrix}$$

De matrix $\mathbf{A}$ heeft dus n rijen en M kolommen.

Elke matrix kan geschreven worden als haar *gereduceerde singuliere waarden ontbinding*:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}}$$

$\mathbf{U}$ heeft n rijen en M kolommen. We noteren de kolommen met $\mathbf{u}_i$. $\mathbf{\Sigma}$ is een diagonaalmatrix met M rijen en M kolommen; de elementen op de hoofddiagonaal noemen we de *singuliere waarden* en noteren we met σ_{ii}. Ze staan bij conventie steeds in dalende volgorde: $\sigma_{11} \geq \sigma_{22} \geq \ldots$. $\mathbf{V}$ heeft M rijen en M kolommen. We noteren de kolommen met $\mathbf{v}_i$ (dus de rijen van $\mathbf{V}^{\mathrm{T}}$ zijn $\mathbf{v}_i^{\mathrm{T}}$). Omdat $\mathbf{\Sigma}$ zoveel nullen heeft, kunnen we ook schrijven:

$$\mathbf{A} = \sigma_{11}\mathbf{u}_1\mathbf{v}_1^{\mathrm{T}} + \sigma_{22}\mathbf{u}_2\mathbf{v}_2^{\mathrm{T}} + \cdots$$

De vectoren $\mathbf{u}_1, \mathbf{u}_2, \ldots$ (elk een n-dimensionale vector, dus interpreteerbaar als een foto) noemt men de *eigengezichten*. Ze beschrijven belangrijkste verbanden tussen de pixels. De singuliere waarden geven weer hoe belangrijk het bijhorende eigengezicht is om de verschillen tussen de foto's te vatten. Daarom nemen we de K eigengezichten met de grootste singuliere waarden. In Figuur 1 zie je de eigengezichten bij de grootste 9 singuliere waarden.

De matrices $\mathbf{U}$, $\mathbf{\Sigma}$ en $\mathbf{V}$ kan je berekenen met het matlabcommando $[U, S, V] = \mathrm{svd}(A, 0)$. Nadat je de K eerste kolommen van $\mathbf{U}$ hebt gecopiëerd, heb je $\mathbf{\Sigma}$ en $\mathbf{V}$ niet meer nodig.

Door de bijzondere eigenschappen van $\mathbf{U}$ kan je de coördinaten van een foto $\mathbf{a}$ berekenen als:

$$c_i = \mathbf{u}_i^{\mathrm{T}}\mathbf{a} \text{ voor } i = 1 \ldots K$$

en zo kan je de afstand tussen foto's berekenen.

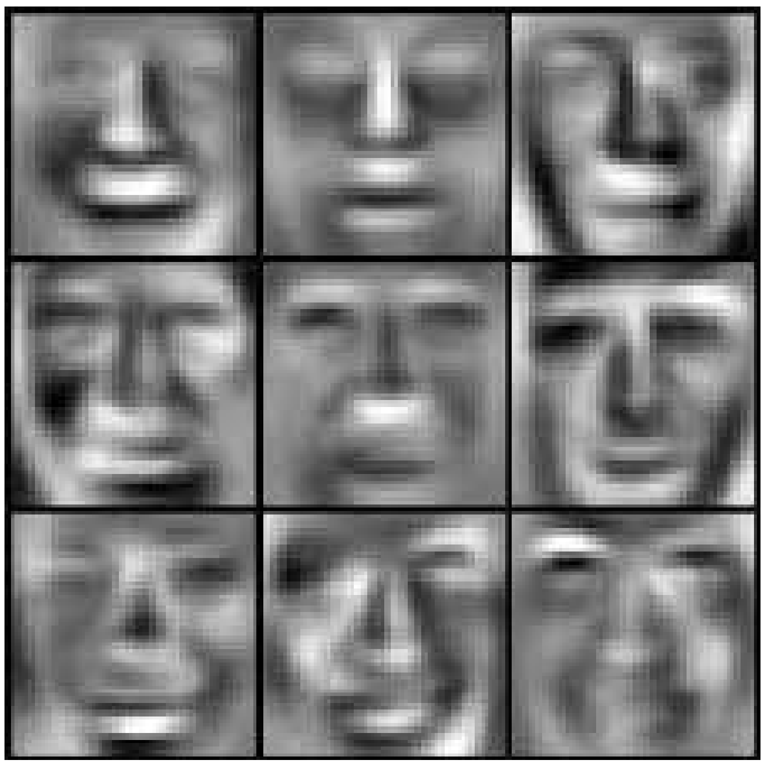
4 Verdieping

Voor zij die reeds over *eigenwaarden* leerden, vermelden we nog het verband tussen eigenwaarden en singuliere waarden. Ter herhaling: een vierkante matrix $\mathbf{Q}$ heeft eigenwaarde λ (een getal) en eigenvector $\mathbf{x}$ als en slechts als $\mathbf{Q}\mathbf{x} = \lambda\mathbf{x}$.

Noteer nu $\mathbf{B} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$ en $\mathbf{C} = \mathbf{A}^{\mathrm{T}}\mathbf{A}$. Dan zijn de $\mathbf{u}_i$ van hierboven de eigenvectoren van $\mathbf{B}$, en de $\mathbf{v}_i$ zijn de eigenvectoren van $\mathbf{C}$. De getallen $(\sigma_{ii})^2$ zijn de bijhorende eigenwaarden (zowel van $\mathbf{B}$ als van $\mathbf{C}$). Bovendien gelden nog de volgende verbanden:

$$\mathbf{u}_i = \frac{1}{\sigma_{ii}}\mathbf{A}\mathbf{v}_i$$

$$\mathbf{v}_i = \frac{1}{\sigma_{ii}}\mathbf{A}^{\mathrm{T}}\mathbf{u}_i$$



Figuur 1: De eigengezichten bij de grootste 9 singuliere waarden.