

Reconhecimento de Faces Utilizando Decomposição em Valores Singulares

Pedro Garcia Freitas Aletéia P. F. de Araújo Mylène C.Q. Farias João Paulo de Andrade Conti
Departamento de Ciências da Computação,
Universidade de Brasília (UnB),
Brasília, Brasil

Resumo—Uma das mais importantes fatorações de matrizes é a Decomposição em Valores Singulares (DVS). Ela é utilizada em diversos problemas práticos, tais como processamento de sinais, ajuste de funções multivariadas, soluções de problemas de otimização, etc. Sua principal aplicação nesses problemas está em permitir a aproximação da pseudo-inversa da matriz decomposta. Neste artigo é apresentado a decomposição DVS e uma amostra de como ela pode ser utilizada para caracterizar informações em imagem de forma que permita a identificação de faces semelhantes.

Index Terms—Decomposição em Valores Singulares, Detecção de Faces, Espaço de Faces, DVS

I. INTRODUÇÃO

Reconhecimento de faces por computador é um importante campo de estudo na área de visão computacional, e tem sido objeto de estudo por vários anos devido à sua aplicação prática na pesquisa e indústria. Por exemplo, controle de acesso à informação baseado em reconhecimento de faces, interação homem-máquina, computação ubíqua e programas com aplicações policiais tem utilizado vastamente esse recurso.

Diversas técnicas computacionais costumam ser combinadas para resolução desse problema. Redes neurais, aprendizado de máquina e estatística computacional são exemplos de recursos complexos que tem sido utilizados para uma solução satisfatória. A maioria dos trabalhos bem sucedidos detectam individualmente características do rosto, tais como distâncias entre olhos, tamanho do nariz, boca, etc. Assim, diversos recursos de morfologia são também utilizados.

Uma abordagem bem-sucedida na literatura para detecção de faces utiliza Análise de Componentes Principais (PCA) para reconhecimento com base em projeção sobre os autovalores de matrizes de correlação. Psarrou et al.[1] mostram como detectar faces projetando-se a imagem a ser reconhecida em um *espaço de faces*, realizando a classificação por comparação com coordenadas das faces conhecidas.

Neste contexto, a ideia do presente trabalho é modificar essa abordagem para utilizar *Decomposição em Valores Singulares* (DVS), a fim de simplificar o modelo matemático e reduzir o número de operações executadas para solução do problema. Métodos de detecção de faces que utilizam projeções sobre o espaço DVS á foram estudados por Chou-Hao e Chen[2], Zhang et al.[3], Nain et al.[4], entre outros. Esses trabalhos possuem a característica em comum de combinarem técnicas de redução de dimensionalidade, seja através de Análise de Discriminantes Lineares, seja

através de PCA ou de outra técnica. O presente trabalho propõe a utilização restrita da decomposição DVS para o reconhecimento facial, o que implica em características de simplicidade algorítmica e independência de dados, o que contribui para implementações com recursos de computação paralela.

A Decomposição em Valores Singulares é um tópico de álgebra linear, sendo componente de diversos métodos numéricos e computacionais, tais como técnicas de resolução de sistemas lineares, aplicações em geofísica[5], genética[6], teoria da informação[7], etc. Em processamento de imagens, há uma vasta utilização desta técnica matemática associada à diversas operações feitas sobre imagens, como, por exemplo, detecção de bordas[8], operações com marcas d'água[9], compressão de imagens[10], etc. Técnicas baseadas em DVS são utilizadas também para caracterização de objetos de imagens, que são aplicações semelhantes à que será proposta neste trabalho, tais como reconhecimento automático de objetos[11] e identificação biométrica digital[12].

II. DECOMPOSIÇÃO EM VALORES SINGULARES

Seja uma matriz A com m linhas n colunas, com posto matricial r tal que $r \leq n \leq m$. Então, A pode ser fatorizada em três matrizes U , S e V . Isto é,

$$A_{mn} = U_{mm} S_{mn} V_{nn}^T \quad (1)$$

onde U e V são matrizes ortogonais de dimensões $m \times m$ e $n \times n$, respectivamente. S é uma matriz ortogonal de ordem $m \times n$ e que possui elementos que caracterizam bem a matriz decomposta, sendo chamada de *matriz de valores singulares*, tendo a forma

$$S_{m,n} = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ 0 & 0 & \cdots & 0 \end{pmatrix} \quad (2)$$

onde os *valores singulares* σ_i sempre satisfazem a propriedade

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_n \geq 0 \quad (3)$$

Além dessa propriedade, tem-se que os *vetores singulares*, $u_i \in U$ e $v_i \in V$ formam sempre dois conjuntos ortonormais. Ou seja, para

$$U = [u_1 \quad u_2 \quad \cdots \quad u_m] \quad (4)$$

tem-se que

$$u_i u_j^T = \delta_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases} \quad (5)$$

e também, de forma análoga,

$$v_i v_j^T = \delta_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases} \quad (6)$$

Quando A é uma matriz simétrica com autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ e seus autovetores associados q_1, q_2, \dots, q_n , pode-se utilizar a *decomposição espectral*[13] de A . Ou seja,

$$A = \sum_{i=1}^n \lambda_i q_i q_i^T \quad (7)$$

Assim, a decomposição pode ser reescrita definindo-se

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (8)$$

e

$$Q = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{bmatrix} \quad (9)$$

e reescrevendo A como

$$A = Q \Lambda Q^T \quad (10)$$

que é uma composição idêntica à DVS quando definimos $U = V = Q$ e $S = \Lambda$.

A partir das definições acima, Ralston e Rabinowitz[14] listam uma série de propriedades do DVS, tais como

- $\|A\| =$ maior autovalor de $\sqrt{A^T A} = \sigma_1$
- $\|A^{-1}\| = \sigma_n \cdot A$
- Os valores singulares são únicos para cada matriz, servindo como uma “identificação” para a informação contida na matriz decomposta.
- Para $v_i \in V$ e $u_i \in U$, temos que v_i e u_i são autovetores de A .
- O posto da matriz A é igual ao número de valores singulares não-nulos, permitindo representar matrizes esparsas de forma comprimida e sem perda de informação.

III. DETECÇÃO DE FACES UTILIZANDO DVS

Seja o conjunto de N imagens armazenadas no banco de dados de faces a serem reconhecidas. Assumindo que todas as imagens deste banco de dados possuem tamanho $m \times n = M$ e são armazenadas após serem convertidas em um vetor coluna de tamanho $M \times 1$. Então, temos que o banco de dados é representado por uma matriz Φ , com N números de faces que formam a matriz $M \times N$

$$\Phi = [\phi_1 \quad \phi_2 \quad \cdots \quad \phi_N] \quad (11)$$

Em seguida, cria-se as colunas da matriz a ser decomposta A

$$a_k = \phi_k - \frac{1}{N} \sum_{i=1}^N \phi_{ki} \quad (12)$$

ficando

$$A = [a_1 \quad a_2 \quad \cdots \quad a_N], \quad (13)$$

que é a matriz de variância de Φ .

A partir de A , utiliza-se a decomposição DVS para gerar as matrizes U , S e V , onde U forma uma base ortogonal.

O próximo passo do algoritmo consiste em decompor cada ϕ_i em projeções de U , gerando o vetor de projeções coordenadas x_i

$$x_i = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} (\phi_i - \bar{\phi}) \quad (14)$$

Assim, seja f o parâmetro de entrada do algoritmo, que é a face a ser reconhecida no banco de dados, já convertida em um vetor de tamanho $M \times 1$, cria-se o vetor de coordenadas x tal que

$$x = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix} (f - \bar{\phi}) \quad (15)$$

O próximo passo consiste em criar o *vetor de classificação*,

$$\Xi = [\xi_1 \quad \xi_2 \quad \cdots \quad \xi_N]^T \quad (16)$$

o qual guarda a distância entre a imagem de entrada e as imagens armazenadas. A ideia é encontrar qual menor elemento de Ξ , a fim de encontrar a menor distância entre as duas. Isto é, para

$$\xi_i = \|x - x_i\| \quad (17)$$

a posição do menor elemento de Ξ equivale à posição da imagem reconhecida no banco. Além disso, utilizamos um limiar ε para determinar uma distância mínima aceitável. Para casos em que todos os elementos $\xi_i < \varepsilon$, f é uma imagem não-identificada. Os passos de identificação estão resumidos no Algoritmo 1.

Algorithm 1 Reconhecimento de faces utilizando DVS

Require: Imagem de entrada f e banco de dados Φ

Ensure: Dizer se f está em Φ

- 1: Computar $\bar{\phi}$
 - 2: Gerar a matriz A .
 - 3: Gerar USV^T realizando a decomposição em valores singulares de A .
 - 4: Computar todos x_i a partir de todos ϕ_i .
 - 5: Computar x a partir de f .
 - 6: Definir ε a distância aceitável para dizer se uma face está no banco.
 - 7: Calcular o vetor de classificação, calculando ξ_i de cada face do banco.
 - 8: Buscar qual menor $\xi_i < \varepsilon$. A imagem é reconhecida como sendo a i -ésima imagem em Φ .
-

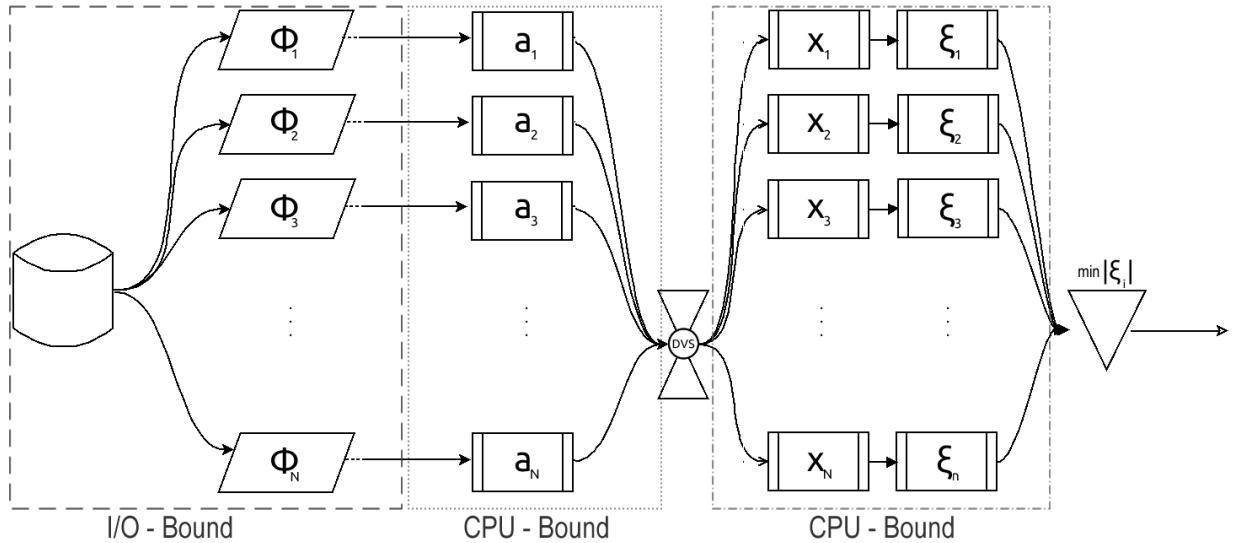


Figura 1. Diagrama de execução paralela.

IV. EXPERIMENTOS

Os testes de reconhecimento foram feitos utilizando-se imagens de faces livres de *copyright* oferecidas pela Universidade da Dinamarca[15], consistindo de um banco de dados com faces frontais, redimensionadas para o mesmo tamanho com 10 indivíduos distintos e todas convertidas para um único canal em escala de cinza. Estas imagens estão apresentadas na Fig. 2.

A primeira parte do experimento constituiu-se de buscar um indivíduo com face armazenada no banco. Para isso, a entrada f foi tomada com posições diferentes para a face já registrada. As imagens escolhidas foram aquelas listadas na Fig. 2. Assim, a partir das quatro entradas escolhidas, o algoritmo reconheceu três. Sendo que para a imagem de semi-perfil foi retornado que a imagem não estava no banco, mostrando que o algoritmo é sensível à posição frontal.

A segunda parte do experimento consistiu na busca de um indivíduo que não estava registrado no banco (Fig. 2). Para esse caso, o algoritmo detectou que a face não teve registro prévio no banco.

V. PARALELIZAÇÃO E APERFEIÇOAMENTOS ALGORÍTMICOS

Na Seção III, discutiu-se os passos para criação do *vetor de classificação* Ξ que nos permite encontrar imagens próximas no banco Φ . Para isso, deve-se comparar o vetor de coordenadas da imagem de entrada, x , com o conjunto de vetores de projeções x_i . Pela Equação 14, nota-se que, após calculado a matriz U , são gerados N vetores de projeções x_i , onde cada um é dependente somente da coluna do banco ϕ_i . Além disso, a partir da Equação 15, podemos notar que o vetor de coordenadas x é calculado somente uma vez. Dessas verificações, nota-se uma grande independência dos dados para o cálculo desses vetores, o que permite concluir que o *vetor de classificação* pode ser gerado a partir de um algoritmo que calcula independentemente cada elemento desse vetor, uma vez que cada $\xi_i \in \Xi$ é dependente somente de x e x_i , calculados anteriormente.

Assim, recursos de computação paralela podem ser utilizados para acelerar o cálculo do vetor de classificação, aproveitando-se essa independência dos dados. Tanto paradigmas que utilizam modelos de memória compartilhada quanto distribuídas podem ser utilizados na implementação dessa paralelização.

Para uma implementação paralela do algoritmo é possível se obter 2 níveis de paralelização: uma associada à entrada e saída (*IO-bound*) e outra associada às operações de cálculos aritméticos (*CPU-bound*). A organização das divisões dos dados para execução em paralelo estão esquematizadas na Fig. 1.

VI. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho apresentou-se um único método computacional, conhecido como Decomposição em Valores Singulares para inferir a presença de uma face em um banco de faces previamente gerado. A partir dessa abordagem conseguiu-se uma forma simples e computacionalmente eficiente de realizar a tarefa de reconhecimento sem utilizar técnicas complexas, tais como Redes Bayesianas[16], [17] ou processos estatísticos mais custosos[18], [19].

Os resultados dos testes no reconhecimento de faces mostram que é necessário um aprimoramento no algoritmo para reconhecer posições não simétricas àquelas que estão registradas no banco. Contudo, o algoritmo obteve sucesso para encontrar as faces que tinham variações em relação àquelas registradas (sorriso, ângulo dos olhos, etc).

REFERÊNCIAS

- [1] R. P. A., S. G. B., and M. W. A., "Eigenfaces for recognition," in *In J of Cognitive Neuroscience v3 nr1*, 1991, pp. 71–86.
- [2] C.-H. Hsu and C.-C. Chen, "Svd-based projection for face recognition," in *Electro/Information Technology, 2007 IEEE International Conference on*, may 2007, pp. 600–603.
- [3] D. Zhang, S. Chen, and Z. hua Zhou, "A new face recognition method based on svd perturbation for single example image per person," *APPLIED MATHEMATICS AND COMPUTATION*, vol. 163, pp. 895–907, 2005.

- [4] N. Nain, N. Agarwal, P. Gour, R. P. Talawar, and S. Chandra, "Face recognition using pca and lda with singular value decomposition (svd)," in *IPCV*, H. R. Arabnia, Ed. CSREA Press, 2008, pp. 387–393. [Online]. Available: <http://dblp.uni-trier.de/db/conf/ipcv/ipcv2008.html#NainAGTC08>
- [5] H. P. Moura, "Decomposição em valores singulares aplicada a dados de campo magnético," *Revista Brasileira de Geofísica*, vol. 16, no. 1, 1998.
- [6] O. Alter, P. O. Brown, and D. Botstein, "Singular value decomposition for genome-wide expression data processing and modeling," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 97, no. 18, august 2000.
- [7] J. Miszczak, "Singular values decomposition and matrix reorderings in quantum information theory," *arXiv:1011.1585*. [Online]. Available: <http://arxiv.org/abs/1011.1585>
- [8] A. Tewfik, F. Assaad, and M. Deriche, "Edge detection using spectral estimation techniques," in *Multidimensional Signal Processing Workshop, 1989., Sixth*, sep 1989, pp. 34–35.
- [9] W. Zhao, Z. Zhai, and Z. Wang, "Digital watermarking algorithm based on block singular value decomposition and bp network," in *Control and Decision Conference (CCDC), 2010 Chinese*, may 2010, pp. 2503–2508.
- [10] M. Dixit, Priyatamkumar, and C. Vijaya, "Computational analysis of adaptive singular value decomposition algorithm to 2d and 3d still image compression application," in *Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference on*, oct. 2010, pp. 482–487.
- [11] Jun-Haizhai, X.-Z. Wang, S.-F. Zhang, and J. Li, "View-based 3d object recognition using wavelet multiscale singular-value decomposition and support vector machine," in *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, vol. 3, nov. 2007, pp. 1428–1432.
- [12] R. Radhakrishnan and C. Bauer, "Robust video fingerprints based on subspace embedding," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 31 2008-april 4 2008, pp. 2245–2248.
- [13] W. Rudin, *Functional analysis*, 2nd ed., ser. International Series in Pure and Applied Mathematics. New York: McGraw-Hill Inc., 1991.
- [14] A. Ralston and P. Rabinowitz, *A first course in numerical analysis*, ser. Dover books on mathematics. Dover Publications, 2001. [Online]. Available: <http://books.google.com/books?id=czHV-1bEF10C>
- [15] M. B. Stegmann, "Annotated face, hand, cardiac and meat images," <http://www2.imm.dtu.dk/~aam/>, 1995, [Online; accessed 2-November-2011].
- [16] A. Nefian, "Embedded bayesian networks for face recognition," in *Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on*, vol. 2, 2002, pp. 133–136 vol.2.
- [17] P. de Souza Coelho, C. Esperanca, and A. Fernandes de Oliveira, "Enhancing the bayesian network approach to face detection," in *Computer Science Society, 1999. Proceedings. SCCC '99. XIX International Conference of the Chilean, 1999*, pp. 150–159.
- [18] H. Demirel and G. Anbarjafari, "Improved face recognition system using probability distribution functions extracted from wavelet subbands," in *Computer and Information Sciences, 2009. ISCIS 2009. 24th International Symposium on*, sept. 2009, pp. 94–98.
- [19] K. Delac, M. Grgic, and S. Grgic, "Statistics in face recognition: analyzing probability distributions of pca, ica and lda performance results," in *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on*, sept. 2005, pp. 289–294.

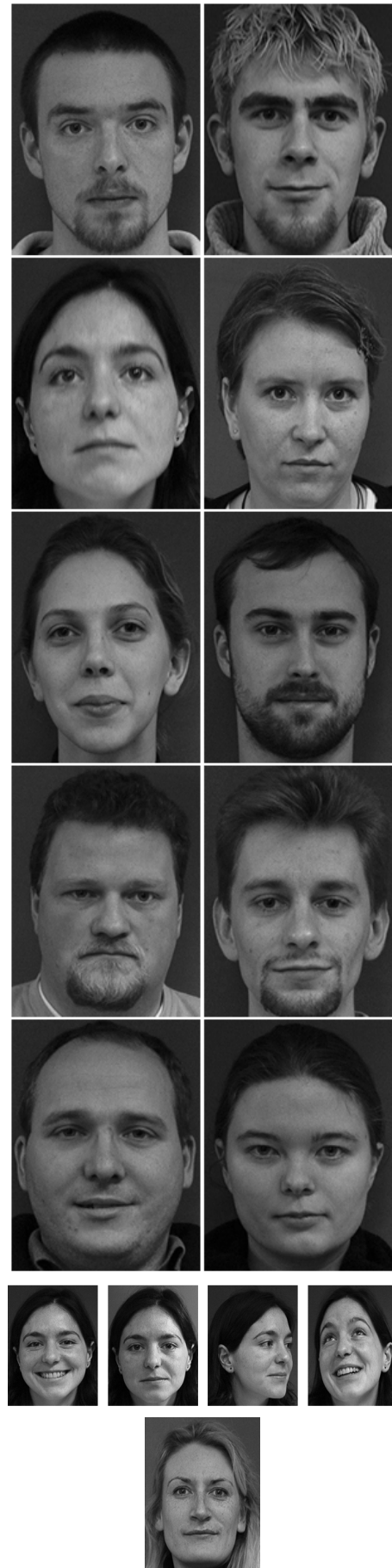


Figura 2. Banco de Imagens (Cima), faces para reconhecimento (meio), Face para reconhecimento não presente no banco (baixo).