Expression Transfer with Multilinear AAM's

Instituto Nacional de Matemática Pura e Aplicada – IMPA
Vision and Graphics Laboratory – Visgraf

Emilio Ashton Vital Brazil `<emilio@visgraf.impa.br>`
Ives José de Albuquerque Macêdo Júnior `<ijamj@visgraf.impa.br>`

Summary

- **Introduction**
- **Active Appearance Models**
- **Multilinear Analysis**
- **Expression Transfer**
- **Results and Discussions**
- **Conclusions**
- **Future Work**

Introduction

- ***Expression transfer definition***
  - ***"Given two images from different subjects map the expression performed by the source subject onto the target subject's face"***
- ***Subproblems***
  - ***Face image capture and representation***
  - ***Face region localization***
  - ***Identity and expression modeling and recognition***
  - ***Photo-realistic resynthesis***

Introduction

- **As many subproblems relates to known problems approached by the *computer vision* community, it is natural to search methods in this literature**
- **We build on the (computer vision) work of *Wang and Ahuja* about facial expression decomposition, who based their work on results of:**
  - ***Cootes, Edwards and Taylor* (AAM's)**
  - ***Vasilescu and Terzopoulos* (Multilinear Analysis)**

Active Appearance Models

- ***Cootes*, *Edwards* and *Taylor* (ECCV 1998)**
- **Generative model-based representation and interpretation of images (*shape* and *texture*)**
- **Closely related to other models for computer vision (e.g. *morphable models* and *active blobs*)**
- **Linear in both features (PCA-based)**
- **Sort of "shape enhanced" *Eigenfaces* (*Turk* and *Pentland*, JCN 1991)**

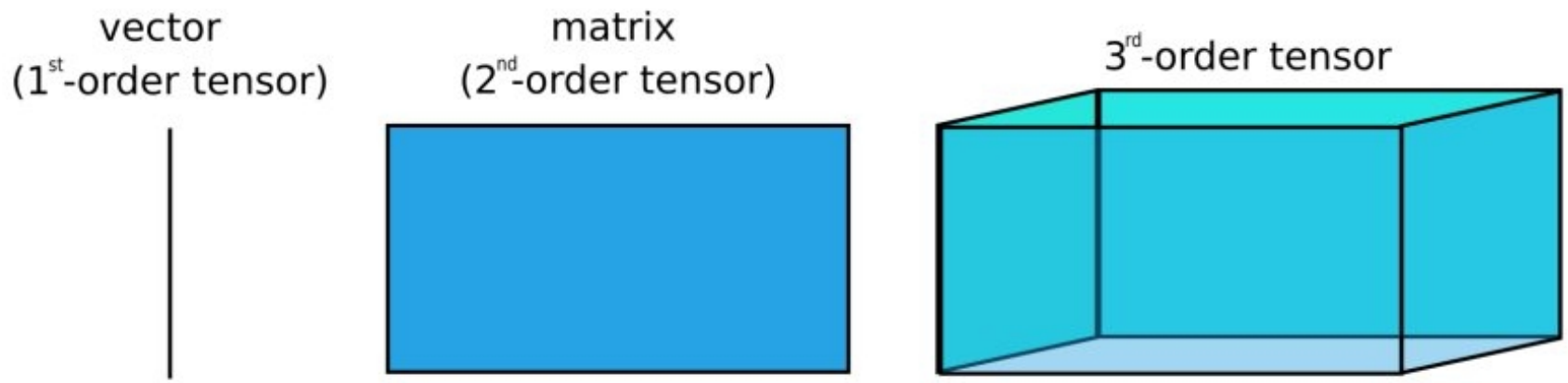
Multilinear Analysis

- **Introduced to the computer vision and graphics communities by *Vasilescu* and *Terzopoulos***
- **Based on *multilinear algebra* (the algebra of higher-order tensors), a generalization of *linear algebra***
- **Designed as a mathematical framework for *analysis* and *representation* of image ensembles**
- **Has as basic object a generalization of vectors and matrices, the *tensor***

Multilinear Analysis

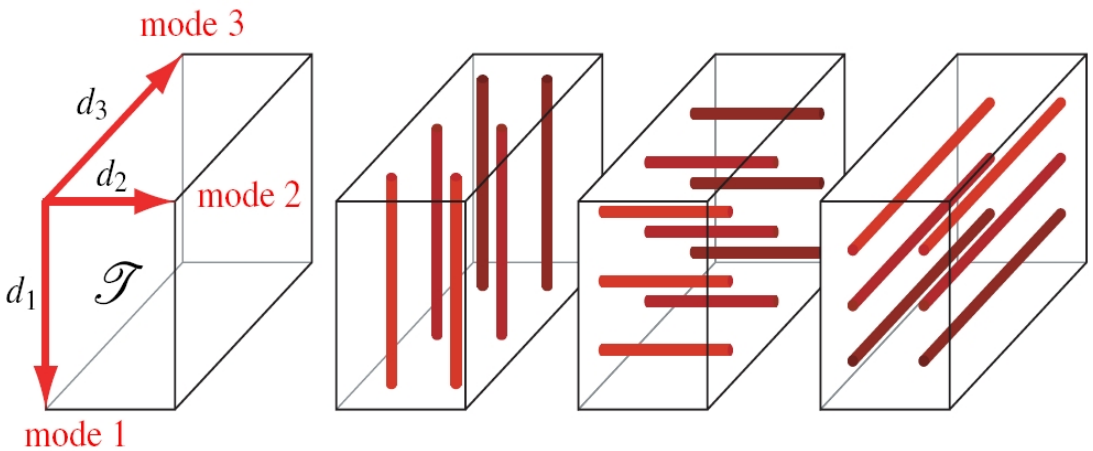
- ***Tensor*: a natural generalization of vectors (*1st*-order tensors) and matrices (*2nd*-order tensors) to multiple indices**

- **an *Nth*-order tensor can be thought of as a block of data indexed by *N* indices, $\mathfrak{T} = (t_{i_1 i_2 \ldots i_N})$**



Multilinear Analysis

- ***Mode spaces*: a matrix has two characteristic spaces, row and column space; a tensor has one for each mode, hence we call then *mode spaces***
  - **below, is depicted a *3*rd-order (or *3*-mode) tensor of dimensions $d_1 \times d_2 \times d_3$**



Multilinear Analysis

- ***Mode*-n *flattening*: the "stacking" of mode-n vectors of a tensor as columns in a matrix, denoted as a subscript (n) after the tensor name: $\mathfrak{T}_{(n)}$ or $\mathbf{T}_{(n)}$**

- ***Mode*-n *product*: a linear transformation on all mode-n vectors of a tensor. It is defined between a tensor $\mathfrak{T}$ and a matrix M and is denoted by $\mathfrak{T} \times_n \mathbf{M}$**

- ***Rank-1 tensor*: is defined by vectors**

  $$\mathbf{u}_1 \in \mathbb{R}^{d_1}, \ldots, \mathbf{u}_N \in \mathbb{R}^{d_N} \quad \textit{as} \quad \mathfrak{T} = \mathbf{u}_1 \otimes \ldots \otimes \mathbf{u}_N,$$

  where $\otimes$ denotes the tensor product

  (note that $\mathfrak{T}$ has dimensions $d_1 \times \ldots \times d_N$)

Multilinear Analysis

- ***Singular Value Decomposition (SVD)*: states that any matrix M can be decomposed as: $\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^{\mathbf{T}}$, in tensor notation it is written as $\mathbf{M} = \Sigma \times_1 \mathbf{U} \times_2 \mathbf{V}$**

- ***N-mode SVD (HOSVD – Higher Order SVD)*: is one (of many possibles) "generalization" of the SVD, denoted as $\mathcal{T} = \mathcal{L} \times_1 \mathbf{U_1} \times_2 \ldots \times_N \mathbf{U}_N$**

  Obs.: unlike in the SVD, the tensor $\mathcal{L}$ (core tensor) isn't a diagonal tensor (this is an example of some properties which don't generalize well to tensors)

Multilinear Analysis

- ***N-mode SVD Algorithm*:**

  (1) For n = 1, ... , *N*, compute matrix $\mathbf{U_n}$ by computing the SVD of the flattened matrix $\mathfrak{T}_{(n)}$ and setting $\mathbf{U_n}$ to be the left matrix of the SVD

  (2) Solve for the core tensor ($\mathscr{L}$) as follows

$$\mathscr{L} = \mathfrak{T} \times_1 \mathbf{U_1}^{-1} \times_2 \ldots \times_n \mathbf{U_n}^{-1} \times_{n+1} \ldots \times_N \mathbf{U_N}^{-1}$$

$$= \mathfrak{T} \times_1 \mathbf{U_1}^{\mathbf{T}} \times_2 \ldots \times_n \mathbf{U_n}^{\mathbf{T}} \times_{n+1} \ldots \times_N \mathbf{U_N}^{\mathbf{T}}$$

Multilinear Analysis

- ***TensorFaces and Multilinear PCA*:**
  - **In a similar spirit as PCA is based on the SVD, *Vasilescu* and *Terzopoulos* generalized the method and developed a *Multilinear PCA***
  - **With this method in hand, they developed the *TensorFaces* representation of image ensembles in a manner analog to the *Eigenfaces* approach (which is based on the linear PCA), allowing for a better control of the multiple varying factors in images**

Expression Transfer

- **Inspired by the *expression transfer* work of *Vlasic, Brand, Pfister and Popovic* (SIGGRAPH 05)**
- **Influenced by the *Identity* $\times$ *Expression recognition* method of *Wang* and *Ahuja* (ICCV 03)**

- **Based on the assumption:**
  "good analysis is the first step to good (re)synthesis"
  Unknown ☺

Expression Transfer

- **Method Overview**
  - ***Data acquisition***
  - ***Training (Linear AAM + Multilinear Analysis)***
  - ***Expression Transfer***

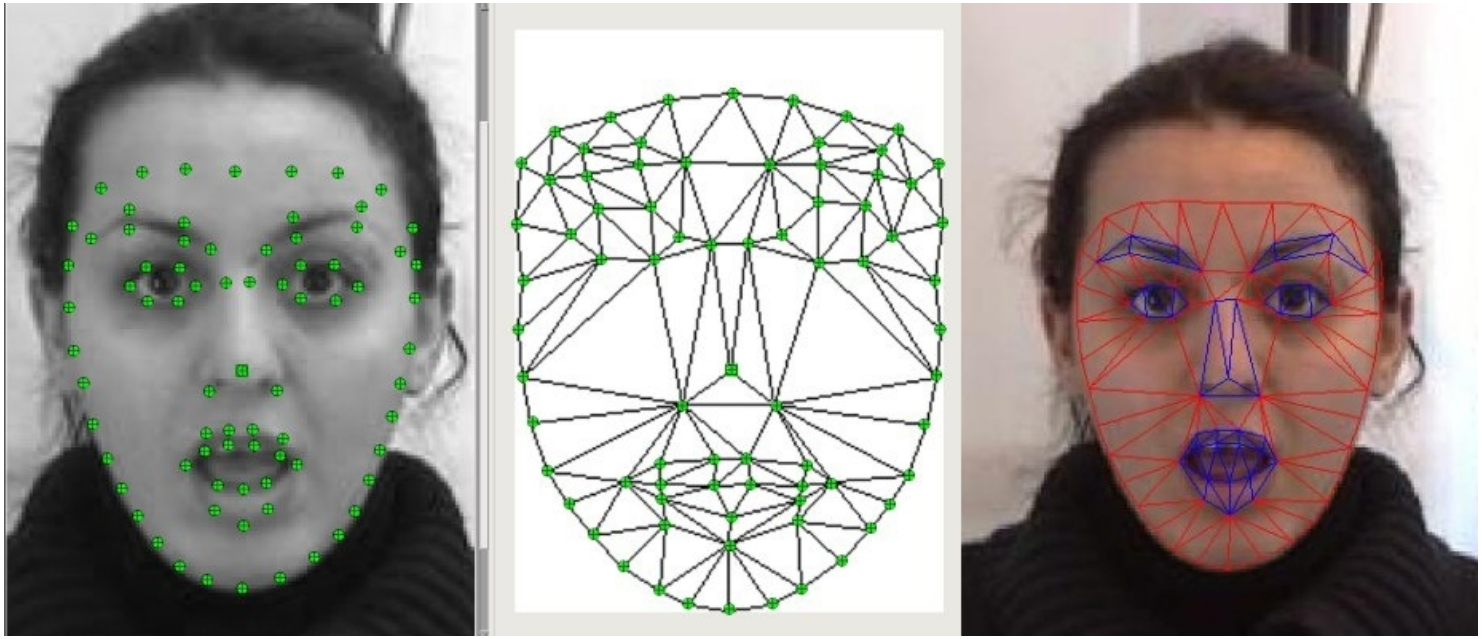
Expression Transfer

- ***Data acquisition – Image capture***
  - **The image data used is from "*The FGnet – Facial Emotion and Expression Database*", kindly provided by *Prof. Frank Wallhof* from *Technische Universität München***
    - **It features *19 subjects* performing the *6 basic expressions* under the same illumination conditions**

Expression Transfer

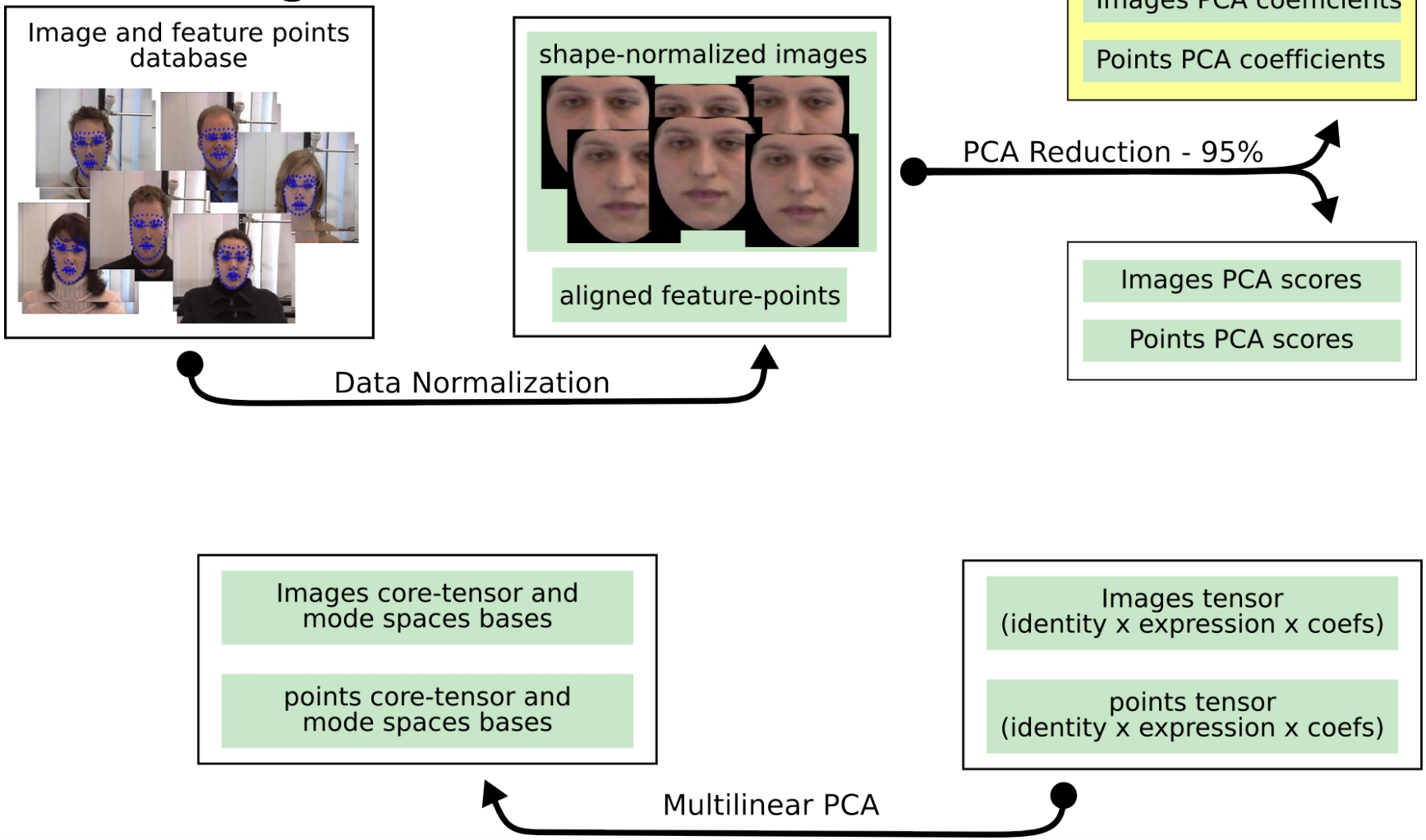
- ***Data acquisition – Feature-Point Marking***
  - ***77*** **feature-points were** ***hand-marked*** **in each of the** ***133*** **images according to a template triangulation of a face**

    (research assistants are welcome !!! ☺)



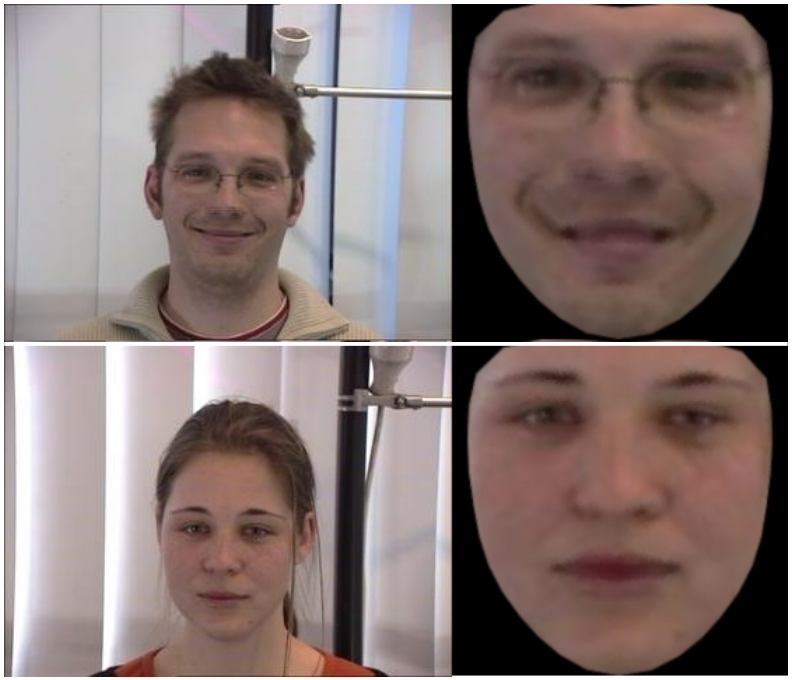
Expression Transfer

- ***Training***



Expression Transfer

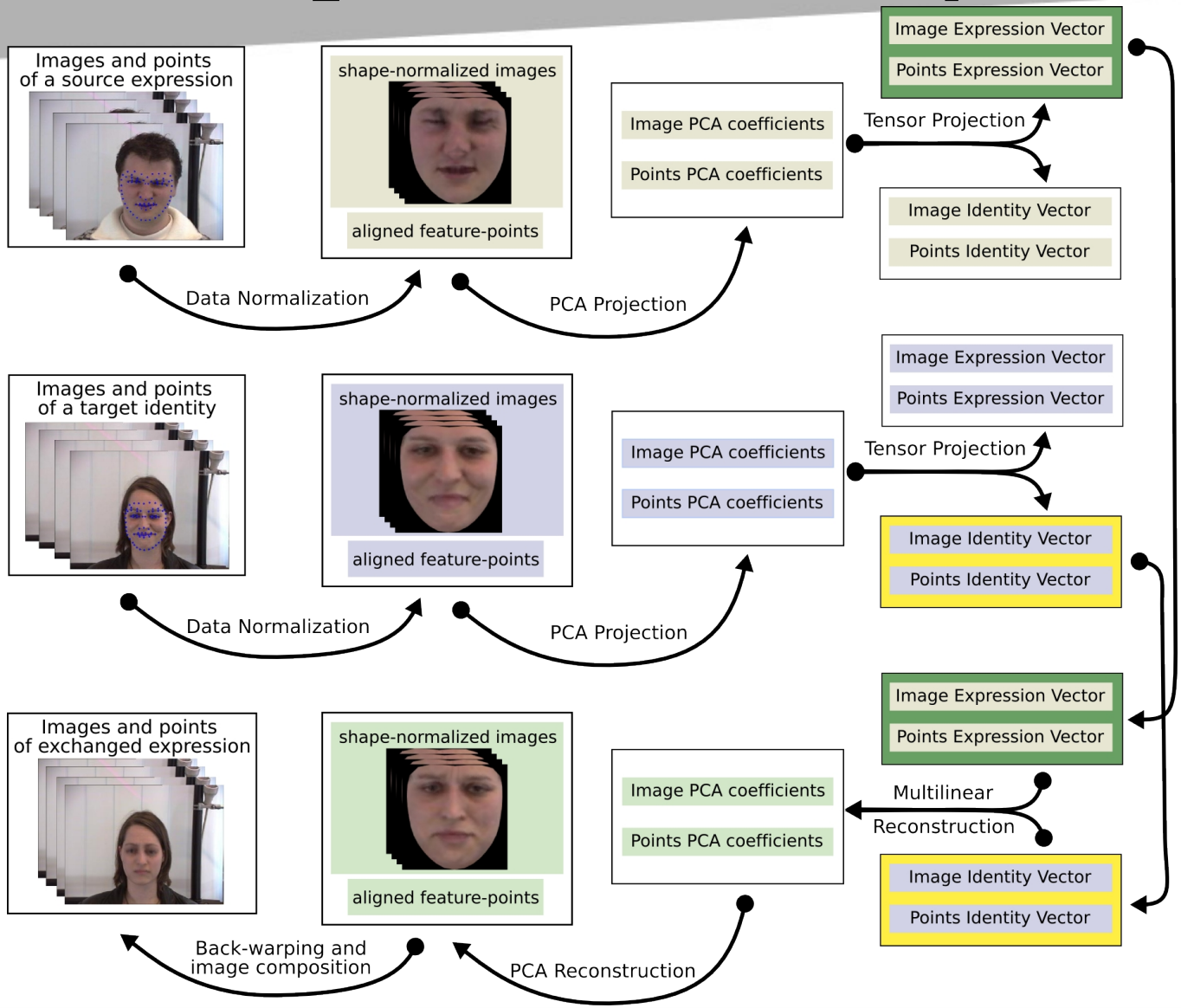
- ***Training – Data Normalization***
  - **Procrustes' alignment of feature-points**
  - **Face extraction and shape-normalization of image (warping)**



Expression Transfer

- ***Training – PCA-based dimensionality reduction***
  - **Separately, we apply a *PCA reduction* (keeping 95% of variance) on the *aligned feature-points* and the *shape-normalized images***
- ***Training – Multilinear PCA (Identity×Expression×Coefs.)***
  - **With the coefficients of our database after PCA reduction, we build two *3rd-order tensors* in which each mode represents *identity*, *expression* and *coefficients variations***
  - **After that, a *Multilinear PCA* is applied to separate the influence along each factor (identity and expression)**

Expression Transfer



Expression Transfer

- ***Expression transfer – Image Normalization***
  - **Feature-points linear conformal alignment**
  - **Shape-normalization of facial image**
- ***Expression transfer – Projection onto PCA basis***
  - **With the aligned feature-points and the shape-normalized texture, the mean is subtracted from each and the result is projected onto the space spanned by the basis acquired after the linear PCA reduction**

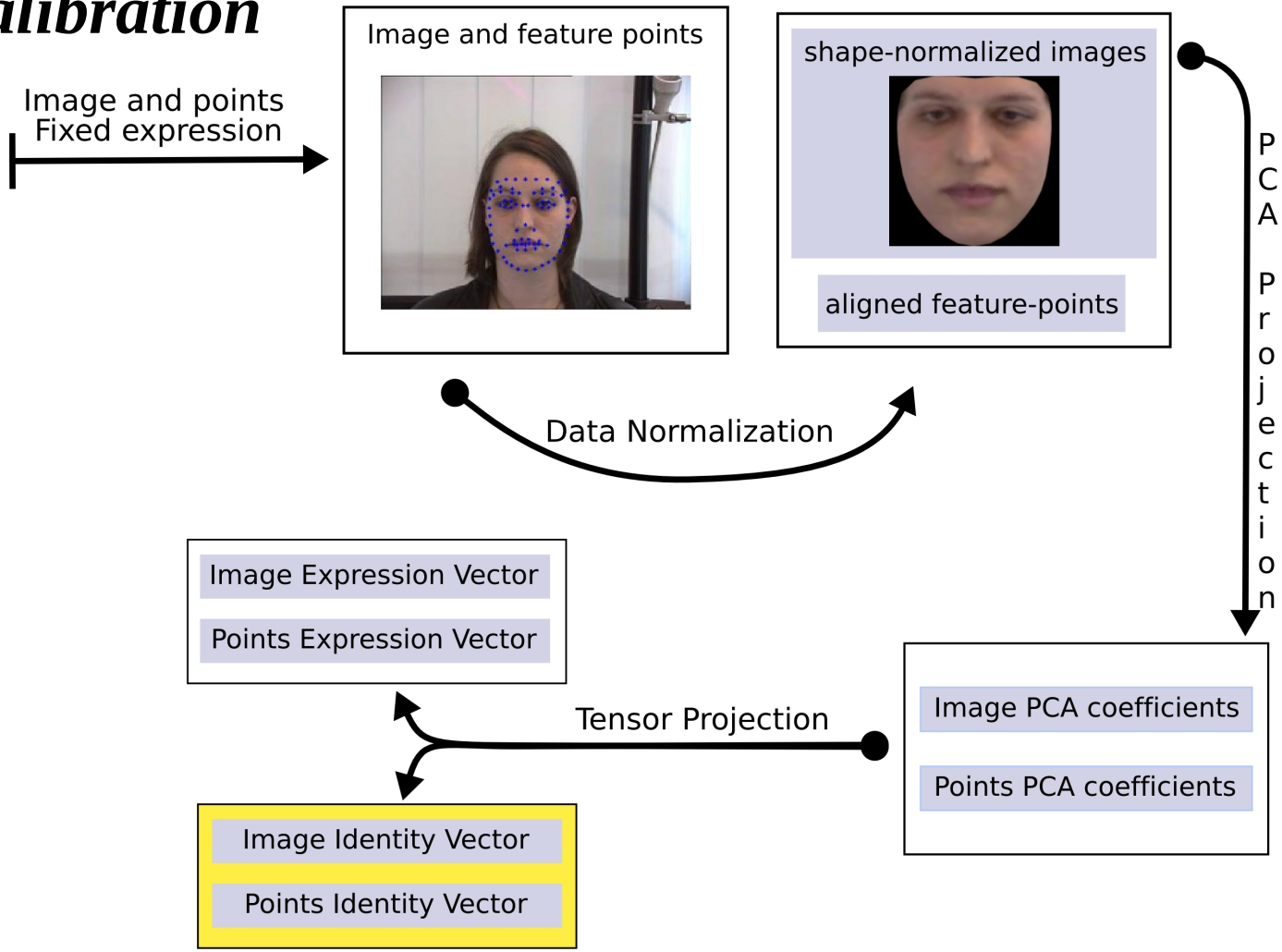
Expression Transfer

- ***Expression transfer – Identity calibration***
  - **Fixating an expression, it's possible to estimate the identity characteristic vector in a given pose via *linear LSQ***

    (as the bilinear tensor becomes a linear operator)
- ***Expression transfer – Expression estimation***
  - **As in the previous, having the identity vector, a new image has the expression vector estimated via *linear LSQ***

Expression Transfer

- ***Identity calibration***

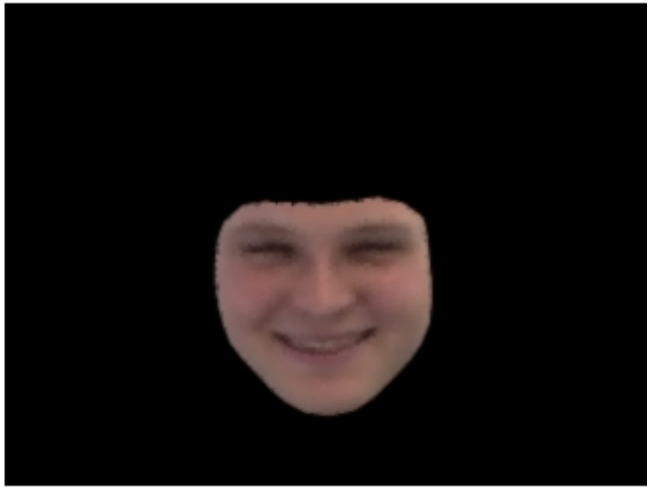


Expression Transfer

- ***Expression transfer – Multilinear reconstruction***
  - **Exchanging the estimated expression vectors, it's possible to reconstruct the new PCA coefficients just applying the tensor on the desired combination of identity and expression vectors**
- ***Expression transfer – PCA reconstruction***
  - **With the new PCA coefficients at hand, a matrix-vector multiplication is sufficient to reconstruct the aligned feature-points and shape-normalized image of the exchanged expression mapped onto the target subject**

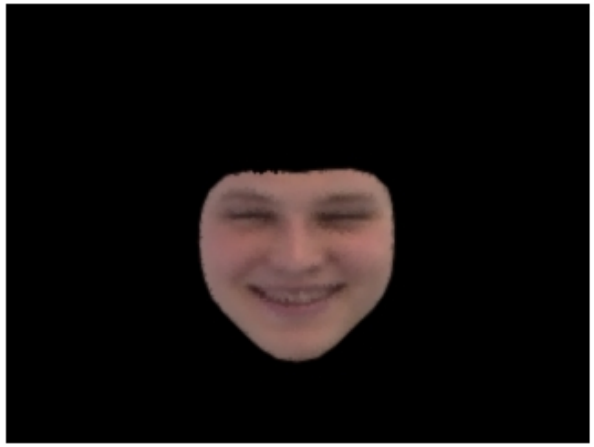
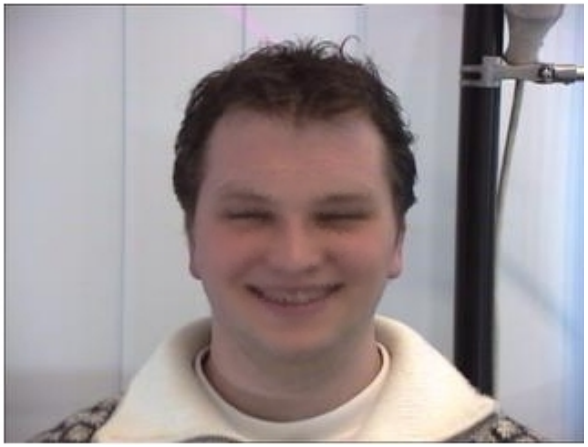
Expression Transfer

- ***Expression transfer – Normalized image (back-)warping***
  - **Applying the inverse alignment transform onto the reconstructed feature-points and mapping the shape-normalized face on their triangulation is possible to warp back the new image**



Expression Transfer

- ***Expression transfer – Face image compositing***
  - **Having the transferred expression mapped onto the target face, an image composition operation is applied and a rewrite of the original image is generated**

Expression Transfer

- ***Expression transfer – The transfer is completed !!!***

Results and Discussions

- **From the original database, 17 (out of 19) subjects where used to train our algorithm and 2 to test it**

Results and Discussions

- **We have tried a number of different approaches and chosen *two* of then as the best with respect to the *trade-off quality/storage***

  (as their computational cost is similar ...)

  – ***{Luminance (from LAB image) / RGB}-based AAM***

  – ***Keeping the variance of the PCA dimensionality reduction on 95%, 98% and 100%***

  – ***Without applying the PCA step and build the tensor directly with aligned feature-points and shape-normalized images***

Results and Discussions

- **We have made *three* different experiments to evaluate the performance (quality) of our approach**
  - ***Projection tests*: where we evaluate how well the original image is represented by our bilinear model**
  - ***Basic expression assignment tests*: where, given an image of a neutral face, we transfer the other 6 basic expressions to it**
  - ***Expression transfer tests*: in which we test the recognition and transfer of basic expression between two different subjects that aren't in the training base**

Results and Discussions

- ***Projection tests (Subject 1 / Luminance)***

Results and Discussions

- ***Projection tests (Subject 1 / RGB)***

Results and Discussions

- ***Projection tests (Subject 2 / Luminance)***

Results and Discussions

- ***Projection tests (Subject 2 / RGB)***

Results and Discussions

- ***Basic expression assignment tests (Subject 1 / Luminance)***

Results and Discussions

- ***Basic expression assignment tests (Subject 1 / RGB)***

Results and Discussions

- ***Basic expression assignment tests (Subject 2 / Luminance)***

Results and Discussions

- ***Basic expression assignment tests (Subject 2 / RGB)***

Results and Discussions

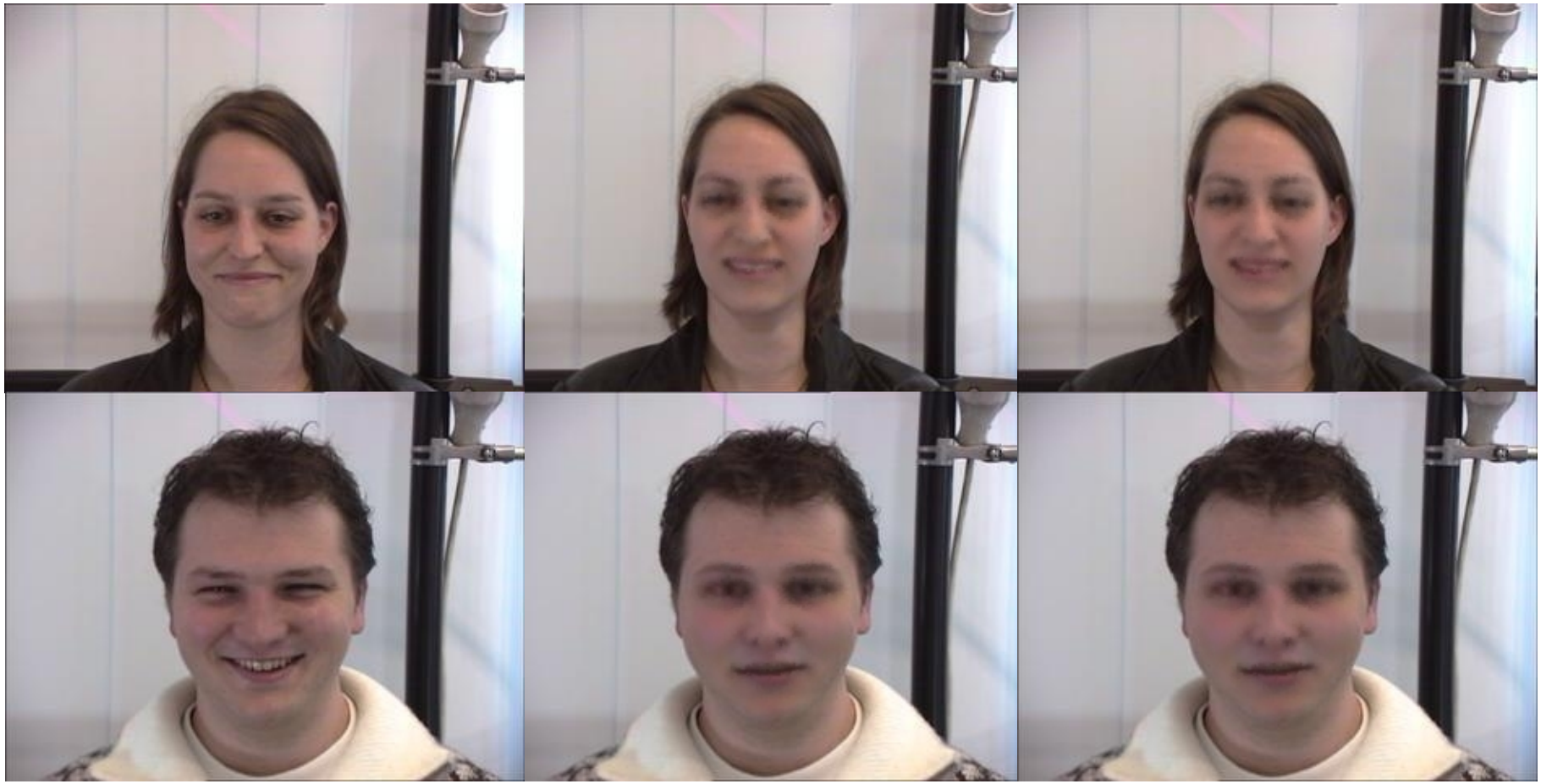
- ***Expression transfer tests***

Results and Discussions

- ***Expression transfer tests***

Results and Discussions

- ***Expression transfer tests***



Results and Discussions

- ***Expression transfer tests***

Conclusions

- **A *luminance* based approach captures rather well photometric details in facial expressions and frees *at least* 2/3 of the total memory requirements !!!**
- **Keeping 95% of the variance, in the PCA reduction, resulted in a good generalization to the modeling of unknown subjects and yielded a good reduction in the total number of PCA coefficients**

Feature-points: 28 (95%), 46 (98%) and 118 (100%)
Normalized images (LAB): 33 (95%), 61 (98%) and 118 (100%)
Normalized images (RGB): 35 (95%), 65 (98%) and 118 (100%)

Conclusions

- ***Facial expression modeling is a hard problem !!!***
- **The use of methods designed primarily for *analysis* is a *good starting point* for *example-based (re)synthesis***
- **Multilinear analysis is a very good approach to model problems involving multiple factors, but ...**
  - **a great amount of data is needed to complete a full tensor, leading to problems in data *acquisition*, *storage* and *management***

Future Work

- **Implement a *model fitting* algorithm to automate the feature-point placement in images (uff !!!)**
- **Experiment with *video rewrite* and *puppetry* (and problems related to the modeling of the complex dynamics of facial expressions)**
- **Evaluate a *rigid-motion invariant representation* for feature-points (eliminate spurious variations from the alignment transformations)**

Future Work

- **Develop a method to allow *incomplete entries* in the tensor *HOSVD* (to deal with acquisition problems and occlusion of some feature-points)**
- **Work on the problem of *illumination inconsistencies* among captured images (a great source of errors)**
- **Experiment with *different architectures* of multilinear separation methods (e.g. a waterfall architecture, where each step solves for a subproblem)**

Future Work

- **Incorporate a new *viseme mode* to our model (comprising a 4th-order tensor) in a manner similar to that proposed by *Vlasic et al*.**
- **Evaluate the quality of *other dimensionality reduction methods* (e.g. *ICA*, *MDS*, *Isomap*, *LLE*, *Laplacian eigenmaps*, *Hessian eigenmaps*, *Principal manifolds*, *Lipschitz embedding*, etc.)**
- **And, at last but not least, generalize theses approaches to model and animate *3D puppets* from monocular video streams**

References

- **T.F. Cootes, G. J. Edwards, and C. J. Taylor. *Active Appearance Models*. ECCV'98.**

- **M. Turk, and A. Pentland. *Eigenfaces for Recognition*. Journal of Cognitive Neuroscience 1991.**

- **M. A. O. Vasilescu, and D. Terzopoulos. *Multilinear Analysis of Image Ensembles: TensorFaces*. ECCV'02.**

References

- **D. Vlasic, M. Brand, H. Pfister, and J. Popovic. *Face Transfer with Multilinear Models*. SIGGRAPH'05.**

- **H. Wang, and N. Ahuja. *Facial Expression Decomposition*. ICCV'03.**

Thank you !!!

Expression Transfer with Multilinear AAM's

Instituto Nacional de Matemática Pura e Aplicada – IMPA

Vision and Graphics Laboratory – Visgraf

Emilio Ashton Vital Brazil `<emilio@visgraf.impa.br>`

Ives José de Albuquerque Macêdo Júnior `<ijamj@visgraf.impa.br>`