

Exploring and Understanding the High Dimensional and Sparse Image Face Space: a Self-Organized Manifold Mapping

Edson C. Kitani¹, Emilio M. Hernandez¹,
Gilson A. Giraldi² and Carlos E. Thomaz³

¹*Universidade de São Paulo, São Paulo, São Paulo,*

²*Laboratório Nacional de Computação Científica, Petrópolis, Rio de Janeiro,*

³*Centro Universitário da FEI, São Bernardo do Campo, São Paulo, Brazil*

1. Introduction

Face recognition has motivated several research studies in the last years owing not only to its applicability and multidisciplinary inherent characteristics, but also to its important role in human relationship. Despite extensive studies on face recognition, a number of related problems has still remained challenging in this research topic. It is well known that humans can overcome any computer program in the task of face recognition when artefacts are present such as changes in pose, illumination, occlusion, aging and etc. For instance, young children can robustly identify their parents, friends and common social groups without any previous explicit teaching or learning.

Some recent research in Neuroscience (Kandel et al., 2000; Bakker et al., 2008) has shown that there is some new information about how humans deal with such high dimensional and sparse visual recognition task, indicating that the brain does not memorize all details of the visual stimuli (images) to perform face recognition (Brady et al., 2008). Instead, our associative memory tends to work essentially on the most expressive information (Bakker et al., 2008; Oja, 1982). In fact, theoretical models (Treves and Rolls, 1994; O'Reilly and Rudy, 2001; Norman and O'Reilly, 2003) have indicated that the ability of our memory relies on the capability of orthogonalizing (pattern separation) and completing (pattern prototyping) partial patterns in order to encode, store and recall information (O'Reilly and McClelland, 1994; Kuhl et al., 2010). Therefore, subspace learning techniques have a close biological inspiration and reasonability in terms of computational methods to possibly exploring and understanding the human behaviour of recognizing faces.

The aim of this chapter is to study the non-supervised subspace learning called Self-Organizing Map (SOM) (Kohonen, 1982; Kohonen, 1990) based on the principle of prototyping face image observations. Our idea with this study is not only to seek a low dimensional Euclidean embedding subspace of a set of face samples that describes the intrinsic similarities of the data (Kitani et al., 2006; Giraldi et al., 2008; Thomaz et al., 2009; Kitani et al., 2010), but also to explore an alternative mapping representation based on manifold models topologically constrained.

More specifically, the purpose of this work is to navigate on the locally optimal pathways composed of the SOM neurons to minimize inappropriate mappings where the standard SOM might show significant discontinuities and compare such visualization procedures on the original image space to understand the most important information captured by the non-supervised model. To minimize image variations that are not necessarily related to differences between the faces, we will carry out experiments on frontal face images available from two distinct public face databases that have been previously aligned using affine transformations and the directions of the eyes as a measure of reference. In this way, the pixel-wise features extracted from the images correspond roughly to the same location across all subjects. In addition, in order to reduce the surrounding illumination and some image artefacts due to distinct hairstyle and adornments, all the frontal images have been cropped to the size of 193x162 pixels, had their histograms equalized and have been converted to 8-bit gray scale. Our experimental results on the two distinct face image sets show that although the standard SOM can explain the general information extracted by its neurons, its intrinsic self-organized manifolds can be better described by an algorithm based on the principle of the locally optimal pathways and the idea of navigating on the graphs composed of the standard SOM neurons.

The remaining of this chapter is organized as follows. In the next section, we briefly review some literature about perceptual and cognitive processes related to human memory and the mechanisms of pattern completion and pattern separation. Next, in the third section, we provide some background definition of SOM and highlight shortly its biological principle of organization that inspired Kohonen in the early eighty's. Also, in the same section, we introduce the standard SOM algorithm based on the competitive learning rule. The main contribution of the chapter is then presented in the subsequent subsection entitled A Self-Organized Manifold Mapping (SOMM) Algorithm. In this subsection, we describe a new algorithm that is able to understand the information extracted from the data, identifying and explaining the nature of the groups or clusters defined by the SOM manifolds. The two distinct public face databases used to carry out the experiments are described in the fourth section. Next, in the fifth section, we show several experimental results to demonstrate the effectiveness of the SOMM algorithm on providing an intuitive explanation of the topologically constrained manifolds modelled by SOM in well-framed face image analysis. Finally, in the last section of the chapter, we conclude this work, summarizing its main points.

2. Neurological and psychological aspects

Several perceptual and cognitive processes guide the task of face recognition in humans. However, one of the most important processes is the memory. Humans do not memorize all the details and features received by the sensory system (Purves et al., 2001). In fact, the human brain has an outstanding capability of forgetting useless information (Brady et al., 2008, Purves et al., 2001).

Basically, human memory can be divided into two groups: declarative and non-declarative memory (Purves et al., 2001). Declarative memory is related to memorizing facts and events and can be accessed for conscious recollection. Facts are information learned during a high level cognition process, such as studying some specific subject. Events are information that one has had as a life experience, for example: birthday, wedding, etc. Episodes at non-declarative memories, on the other hand, are information that cannot be accessed formally.

In other words, it cannot be explained explicitly by words and neither how it occurs nor happens. Examples of non-declarative memory are: physical skills such as swimming, riding a bicycle, or emotional responses such as fear, happiness, etc. Additionally, memories are also categorized as short-term-memory and long-term-memory (Purves et al., 2001). Short-term memories have a limited capacity to hold information and consequently retain it during short period of time (Anderson; 2005), but long-term ones tend to retain it permanently. The process that converts information into long-term memory is known as memory consolidation (Bear, Connors, Paradiso; 2007). The memory consolidation is part of our learning process and is strongly necessary, for instance, to the face-matching task (Kandel et al., 2000).

The brain area responsible for storing the declarative memory is called the Medial Temporal Lobe (MTL) (Bear, Connors, Paradiso; 2007). The MTL is a complex interconnected systems of the brain and one of its most important structures is the hippocampus. Recent experiments carried out on rats have showed that lesions at the hippocampus might affect our capability of learning and retaining information (Bear, Connors, Paradiso; 2007). Yet, in the past, a computational model presented by Treves & Rolls (Treves & Rolls, 1994) had already indicated that some parts of the hippocampus seem to create a sparse and orthogonalized representation of our sensory input and episodic memories. Currently, there is no doubt that the hippocampus plays an important role to encode new episodic memories and, additionally, to prevent the risk of forgetting past memories (Kandel et al., 2000, Kuhl et al., 2010).

Using high-resolution (1.5 millimeters isotropic voxels) functional Magnetic Resonance Imaging (fMRI), Bakker et al. (Bakker et al., 2008) have studied the activity in the human brain MTL area on a set of pattern visualization experiments. The experiments consisted of presenting to each one of a total of eighteen volunteers a sequence of pictures of common objects, such as apples, toys duck, thread balls, wall outlet and etc. The set of pictures used is composed of 144 subsets of slightly different images of the same object, with essentially variation in pose and rotation. The authors have noticed that several brain structures of the MTL area, especially a specific area of hippocampus named CA1, have been activated when pictures of the same object have been presented repetitively and in a interleaved way.

In fact, our brain process of retrieving information can be further described by two main mechanisms: pattern completion and pattern separation (Kuhl et al., 2010). The mechanism of pattern completion is essentially related to the problem where the incoming pattern of some sensory input and the pattern stored in the memory are not exactly the same, but share some similarities. In the mechanism of pattern separation, the similarities between the incoming and stored patterns, if do exist, are minimal and both patterns have, in contrast, a strong degree of dissimilarities that can be mathematically considered as non-correlated or orthogonal.

This work focuses on the mechanism of pattern completion and the role of the human brain hippocampus as an associative memory to propose a new algorithm for the SOM competitive neural network proposed by Kohonen (Kohonen; 1982). Since this pioneering work, it has been argued that SOM is not only a computational approach for data mining and clustering, but also a credible framework at the functional and neural levels to create a self-organization of the input space (Rolls; 2007) and model the human memory activities of encoding and retrieving information.

3. Self-Organizing Map (SOM)

A formal definition of organization is quite complex because it depends on the context. Some crystal structures are considered highly organized due to their symmetry and structural repetition. Functions and hierarchy organize all biological structures, such as the nervous system, digestive system, circulatory system, etc (Kandel et al., 2000). However, in both cases, the definitions of "organization" are ambiguous.

For crystal structures, one finds symmetries and redundancy; on the other hand, a biological system is organized by functions. However, both definitions have in common the sense of similarity that allows us to cluster and hierarchize input patterns. In other words, organization is an association and composition of parts to explore a whole structure or behavior (Asby, 1962; Atlan, 1974).

According to the definition above, clustering is quite related with similarities or even dissimilarities. SOM is an unsupervised neural network developed by Kohonen (Kohonen, 1982; Kohonen, 1990) based on the biological principle of somatosensory organization. According to Kandel et al. (Kandel et al., 2000), there is a functional organization of perception and movement in human and mammals brain. There is also a specialized area in the brain cortex that organizes information coming from sensory pathways or going to motor control. Somatosensory cortex is the area accounting for organizing stimulus coming from different sensory systems, grouping them according to their similarities. In a similar fashion, motor cortex has surfaces dedicated to controlling parts of the body related to movement. This organization in substructures by functions is well-known by neuroscientists, however, why the brain creates this organization remains unclear (Purves et al., 2001).

Based on the biological principle of organization, Kohonen postulates that there are some reasons to have this organization: a) grouping similar stimulus minimizes neural wiring; b) creates a robust and logical structure in the brain, avoiding "crosstalk"; c) from information organized by attributes a natural manifold structure from input patterns can emerge; and d) reduces dimensionality by creating representations (codebooks vectors) that preserves neighborhood relationship between input patterns. Each codebook, also known as BMU (Best Match Unit), retains the most important invariant features that represent a group of input patterns, characterizing an arguable but intuitively analogous behaviour to the pattern completion mechanism of the human brain.

3.1 The Standard SOM algorithm

SOM can be defined as an unsupervised artificial neural network that maps a nonlinear relationship between input patterns in high dimensional space and makes this relationship an ordered and smoothed mapping of input data manifold. SOM has a competitive learning rule, but does not have a rule of convergence or function to minimize. Instead, the algorithm of SOM works with a number of interactions during weight adaptation.

Figure 1 illustrates a Kohonen network of 3×3 output neurons fully connected to the input layer composed by only two neurons. The network is created from a 2D lattice of 'nodes' composed of the output neurons and the input layer. Each output neuron has a specific position $(x, y) \in \mathbb{R}^2$ and contains a vector of weights of the same dimension as the input vector. That is, if the network has m output neurons and the training set consists of vectors $(x_0(t), x_1(t), x_2(t), \dots, x_{n-1}(t)) \in \mathbb{R}^n$, then we have $m \times n$ weights $w_{ij}(t)$, $0 \leq i \leq M-1$, $0 \leq j \leq n-1$ to set.

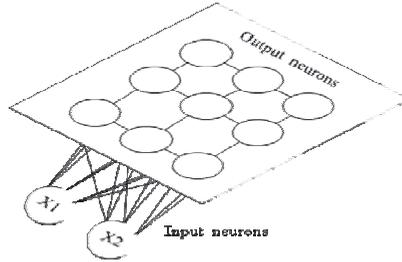


Fig. 1. An illustration of Kohonen network. Note that all input neurons are connected to all output nodes.

The algorithm can be described as follows:

1. Initialize network,
 - 1.1) Define the number m of output neurons that will compose the map and their lattice position (nodes): $r_0, r_1, r_2, \dots, r_{m-1} \in \mathbb{R}^2$,
 - 1.2) Define $w_{ij}(t), 0 \leq i \leq M-1, 0 \leq j \leq m-1$, to be the weight from input neuron i to output neuron j at time t , where M is the size of the set of input training patterns. Initialize weights to small random values. Set the initial radius of neighbourhood around node j , denoted by $\sigma_j(0)$ to be large,
 - 1.3) Define the number of iteration $T \gg M$,
2. Present the input vector $(x_0(t), x_1(t), x_2(t), \dots, x_{n-1}(t)) \in \mathbb{R}^n$, where $x_i(t)$ is sent to the input node i at the time t , where n is the dimensionality of input space,
3. Compute the distance d_j between the input vector $x_i(t)$ and each output neuron j ,

$$\text{given by } d_j = \min_{0 \leq i \leq M-1} |x_i(t) - w_{ij}(t)|, \quad 0 \leq j \leq m-1,$$

4. Designate the BMU neuron r_c to be one with minimum d_j ,
 Update the weights for node r_c and its neighbors, defined by the neighborhood size $\sigma_c(t)$. New weights will be:
 $w_{ij}(t+1) = w_{ij}(t) + \alpha(t)h_{ci}(t)(x_i(t) - w_{ij}(t))$,
 where $\alpha(t)$ is the learning factor:

$$\alpha(t) = \alpha_0 \left(1 - \frac{t}{T} \right),$$

$h_{ci}(t)$ gives the amount of influence that a neuron r_i has on its learning as a function of its distance from the BMU neuron r_c :

$$h_{ci}(t) = \exp \left(- \frac{\|r_c - r_i\|^2}{2\sigma_c^2(t)} \right),$$

Finally, $\sigma_c(t)$ define the radius of influence of the BMU, which can be computed by:

$$\sigma_c(t) = \sigma_0 \exp \left(- \frac{t}{\lambda_c} \right), \text{ and } \lambda_c \text{ is an integer number related to the time of influence of the neighbor radius,}$$

5. Return to step 2 until $t = T$.

From an initial distribution of random weights, the SOM eventually settles into a map of stable zones after some iterations. The term $\alpha(t)$ is a gain term that decreases in time so

slowing the learning process. Besides, the neighbourhood size $\sigma_c(t)$ decreases in size as time goes on, thus localizing the area of maximum activity.

3.2 A Self-organized manifold mapping algorithm

Several studies have provided us with some insight about how to interpret the output of SOMs (Brugger et al., 2008; Bauer & Pawelzik, 1992; Kiviluoto, 1995). One of the best-known tools in this regard is the U-Matrix (Utsch, 2003) that gives us a quantitative summary of the topological relationships between similar data samples. The result of the U-Matrix map is a complex image (coloured or monochromatic) indicating peaks and valleys that represent Euclidean distances between neighbored neurons.

Essentially, the resulting map preserves the topological distribution at the input space of the entire sample data considered. Figure 2 illustrates an example of a coloured U-Matrix map and its hexagonal 5×4 SOM, where each neuron w_{ij} , $0 \leq i \leq M$, $0 \leq j \leq m-1$, has been arbitrary identified by a number. It is possible to see at least two groups of patterns in blue separated by a central chain in red. The chain of high values in the U-matrix indicated by the reddish colours is a representation of some prototypes that are far from both groups and probably describe some data outliers with distinct information about the dataset considered.

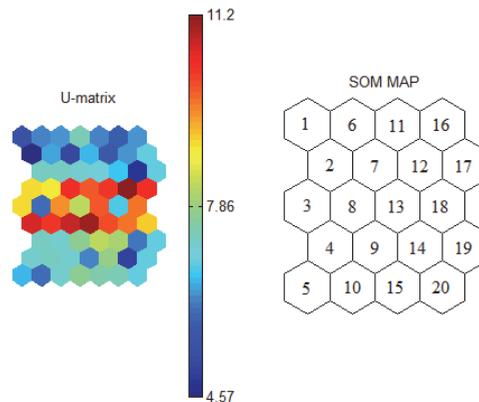


Fig. 2. An illustration of a coloured U-Matrix map and its corresponding SOM where each neuron has been arbitrary identified by a number.

However, to understand the relationship between the information captured in the U-Matrix and the samples, as well as to identify and explain the nature of the groups or clusters defined by the manifolds, it would be helpful to represent all the SOM neurons and their corresponding similarities and dissimilarities on the original data space.

Based on the principle of the locally optimal pathway and the idea of navigating on the neurons that compose the SOM, we propose an algorithm named Self-Organized Manifold Mapping (SOMM) that seeks the pathways or manifolds described by the standard SOM. The SOMM algorithm can be described as follows:

- Calculate the SOM composed of k neurons using the standard Kohonen's algorithm. Create the list $A = \{0, 1, 2, \dots, k-1\}$;
- Calculate pairwise the Euclidean distance d_{ij} of all k neurons;

- c. Create a $(k \times k)$ matrix with all pairwise distance between all k neurons;
- d. Create the list $V = NULL$. Set $r = A_1$, $d_{\min} = \infty$, $i = 0$;
 - d.1) Insert $V \leftarrow \{r\}$,
 - d.1.a) $i = i + 1$.
 - d.2) If $A - V = \emptyset$ go to (e).
 - d.3) Find $d = \min(d_{sr}, s \in A - \{V_{i-1}, V_i\})$. Let s^* such that $d = d_{s^*r}$.
 - d.4) If $s^* \in \{V\}$ go to (e) Else set $r = s^*$ and $V \leftarrow \{r\}$. Go to step (d.1a).
- e. LOOP:
 - e.1) If $V_i = LOOP$ go to (f)
 - e.2) Find $d = \min(d_{sr}, s \in V - \{V_{i-1}, V_i\})$. Let s^* such that $d = d_{s^*r}$.
 - e.3) Insert $V \leftarrow \{s^*\}$, $i = i + 1$.
- f. f) Group BMUs according the order: V_1, V_2, \dots ,
- g. If $A - V \neq \emptyset$ then
 - g.1) $A = (A - V) \cup \{V\}$,
 - g.2) $k =$ number of elements in A,
 - g.3) go to (b).

A simple way to explain this algorithm is to understand the output neurons, represented by the weights w_{ij} computed in the SOM algorithm, as a set of nodes of a fully connected graph (Cormen et al. 2001; Pözlbauer, Rauber, Dittenbach, 2005; Mayer, Rauber, 2010) in the parameter space. Each edge in this graph has a cost given by the Euclidean distance between its ends. Therefore, the $k \times k$ matrix calculated in steps (b)-(c) is a symmetric one holding the edge costs in the graph.

More specifically, in step (d) it is created a list V and in step (d.1) the algorithm inserts in V each visited node. Given a node r , the step (d.3) seeks for the closest neuron s^* such that $s^* \notin V - \{V_{i-1}, V_i\}$; that means, s^* does not belong to the last visited edge of the graph. This step implements a greedy algorithm that makes the locally optimal choice at each stage generating a *locally optimal pathway* that connects a subset of SOM neurons. This is necessary because the idea is to generate a pathway that crosses different clusters but without losing the notion of similarity in the parameter space. If $s^* \in \{V\}$ then we have a loop, like the one exemplified in Figure 3. In this case, the pathway that starts at node 1 ends in the loop $(11) \rightarrow (3) \rightarrow (4) \rightarrow (11)$. The step (e) completes the pathway, which in this figure is composed by the sequence: $V = (1) \rightarrow (2) \rightarrow (11) \rightarrow (3) \rightarrow (4) \rightarrow (11)$.

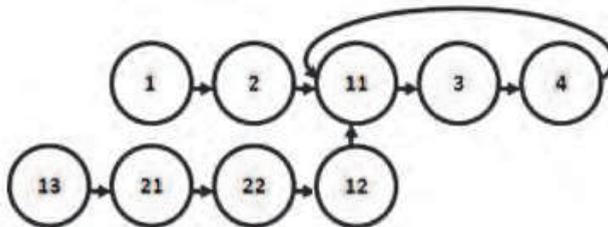


Fig. 3. Two connected pathways with a common loop. The first one starts at node 1 and finally enters in the loop $(11) \rightarrow (3) \rightarrow (4) \rightarrow (11)$, whereas the second path starts in the node 13 and ends in the same loop.

Additionally, the step (g) identifies that there are nodes still not visited by the algorithm. Following the idea of crossing different clusters we must allow that a node $r \in A - V$ might be connected with a node $s \in V$, like node $r = 12$ shown in Figure 3. In terms of the algorithm, it is equivalent to consider V as a node in a new graph (steps (g.1)-(g.2)), compute the new distances in step (b) and seek for another pathway as before. Therefore, this novel algorithm brings the possibility of uncovering clusters not visible by U-Matrix technique or the standard SOM approach.

4. Face databases

We have used frontal images of two distinct face databases publicly available to carry out the experiments. The first database is maintained by the Department of Electrical Engineering of FEI, São Paulo, Brazil (Thomaz and Giraldo, 2010). In this dataset, the number of subjects is equal to 200 (100 men and 100 women) and each subject has two frontal images (one with a neutral or non-smiling expression and the other with a smiling facial expression), so there is a total of 400 images with no significant differences in skin colour to perform the high dimensional and sparse image face analysis. The second dataset is the well-known FERET (Philips et al., 1998) database. In the FERET database, we have considered only 200 subjects (107 men and 93 women) and each subject has two frontal images (one with a neutral or non-smiling expression and the other with a smiling facial expression), providing a total of 400 images with significant differences in skin colour to perform as well the experiments.

To minimize image variations that are not necessarily related to differences between the faces, we previously aligned all the frontal face images using affine transformations and the directions of the eyes as a measure of reference so that the pixel-wise features extracted from the images correspond roughly to the same location across all subjects. Also, in order to reduce the surrounding illumination and some image artefacts owing to distinct hairstyle and adornments, all the frontal images were cropped to the size of 193x162 pixels, had their histograms equalized and were then converted to 8-bit gray scale. Figure 4 illustrates some samples of the FEI (top row) and FERET (bottom row) datasets, highlighting samples of distinct gender, age, facial expression and ethnicity.



Fig. 4. Some samples of the FEI (top row) and FERET (bottom row) frontal images used in the experiments after the pre-processing procedure that aligned, cropped and equalized all the original images to the size of 193x162 pixels.

5. Experimental results

All the experiments have been carried out using the well-known SOM-Toolbox for Matlab created and released by CIS-Helsinki University of Technology (Vesanto, 1999). To address the memory issues related to computing the SOM on high-dimensional datasets, instead of analysing the SOMM algorithm directly on the pre-processed FEI and FERET face images, Principal Component Analysis (PCA) (Fukunaga, 1990) has been applied first to provide dimensionality reduction. However, in order to reproduce the total variability of the sample data, we have composed the PCA transformation matrix by selecting all the principal components with non-zero eigenvalues. Although some of these principal components might represent non-relevant information to understand the differences between the data samples, we are able to represent and further reconstruct the original images without adding any dimensionality reduction artefacts (Kitani et al., 2010).

We have divided our experimental results into two parts. Firstly, we have carried out some face image analyses to understand and visualize the pathways found by the SOMM algorithm where there are subtle differences between the data samples. Thus, we have used a subset of the FEI database composed of non-smiling and smiling face images of females only. Then, in the second part, we have investigated the usefulness of the SOMM algorithm on exploring and understanding the high dimensional and sparse image face space where the differences between the samples are not only related to facial expression but also to gender, ethnicity and age. The goal of the second experiment is to pose an alternative analysis where the differences between the samples are evident, using the whole two FEI and FERET datasets described in the previous section.

Figure 5 illustrates the standard SOM (top left), the pathways described by the SOMM algorithm (bottom left) and their corresponding visualization (top and bottom right) on the original face space using a subset of the FEI database composed of non-smiling and smiling face images of females only. It is important to highlight that since the SOMM navigation is based on the principle of the locally optimal path, it is only possible to visit a new neuron when its distance is minimal regarding all the other neurons previously visited. Therefore, the algorithm explicitly describes the discontinuities present at the high dimensional face image space due to the limited number of input samples. In other words, it is possible to see that SOMM could not find a unique graph that defines a single locally optimal path from non-smiling to smiling female face images. In fact, as shown on the bottom right part of Figure 5, we can see three feasible pathways or clusters: (1) samples that describe a definite smiling facial expression; (2) samples that describe the visual differences from non-convincing to convincing smiling facial expressions; (3) samples that describe the visual differences from non-convincing to convincing non-smiling facial expressions.

In the next two figures, we show the behaviour of the SOMM algorithm on navigating at high dimensional and sparse image face spaces where the differences between the samples are not only related to facial expression but also to gender, ethnicity and age. Figure 6 illustrates the standard SOM (top left), the pathways described by the SOMM algorithm (bottom left) and their corresponding visualization (top and bottom right) on the original face space using the whole set of frontal face images of the FEI database with both gender and facial expression differences.

Analogously to the previous results, three clusters have been found by the SOMM algorithm. Despite the gender differences available on this dataset, SOM has not clearly extracted this information on its standard mapping and neither SOMM has described it in a separated pathway or cluster. The smallest SOMM cluster, composed of 6 neurons, shows samples that describe a definite smiling facial expression with slightly more male facial

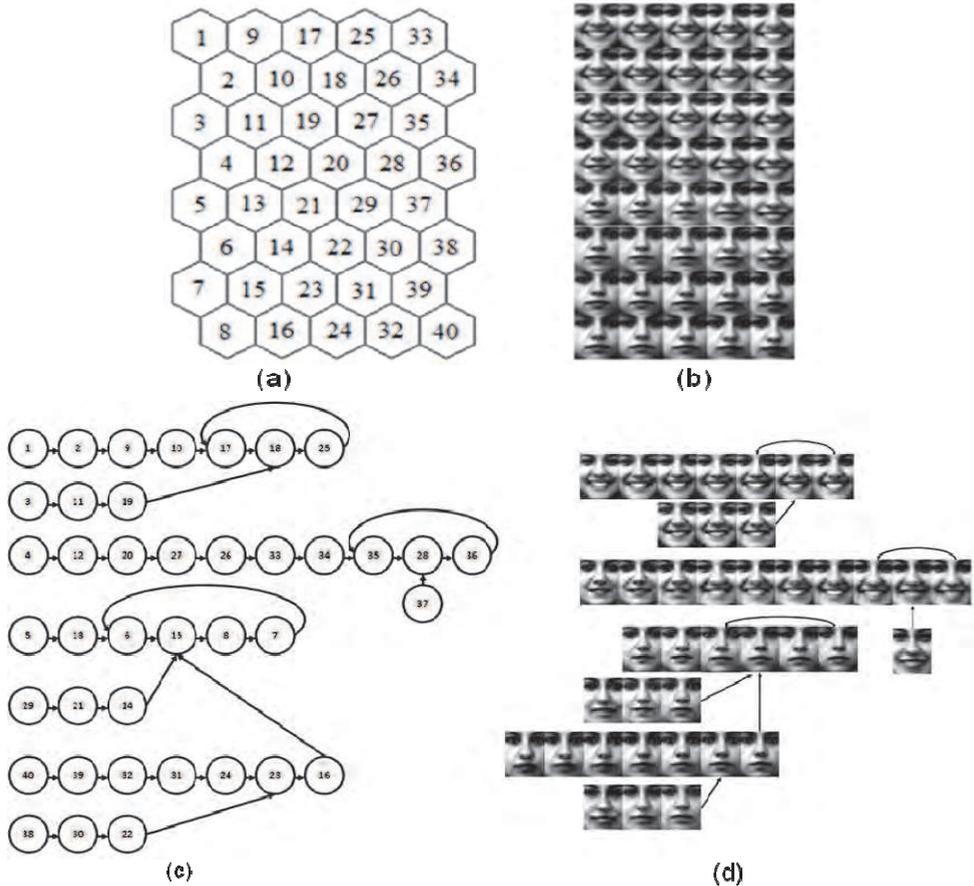


Fig. 5. Exploring the similarities and discontinuities of the high dimensional image face space composed of smiling and non-smiling female face images only of FEI database: standard SOM of size 8x5 (top left); visualization of the SOM neurons (top right); SOMM algorithm navigation (bottom left); visualization of the SOMM clustering (bottom right).

traits than female ones. A similar description is valid for the second smallest SOMM cluster, composed of 8 neurons, but rather with more female facial traits. However, the largest cluster clearly shows that the most expressive information captured by SOMM has been related to changes in facial expression, no matter the gender of the subjects analysed.

The last experimental results using the FERET dataset are presented in Figure 7. It can be seen that the main expressive information captured by SOM have been based on ethnicity and facial expression changes. The visualization of the standard SOM, illustrated on the top right part of Figure 7, shows clearly how the data set has been generally spread along the high dimensional face image space. It is possible to see that when we move from top to bottom we are able to see differences related mainly to ethnicity, no matter the facial expression or gender of the subjects. Besides, navigation on the SOM neurons from left to right highlights essentially information about changes on facial expression with minor

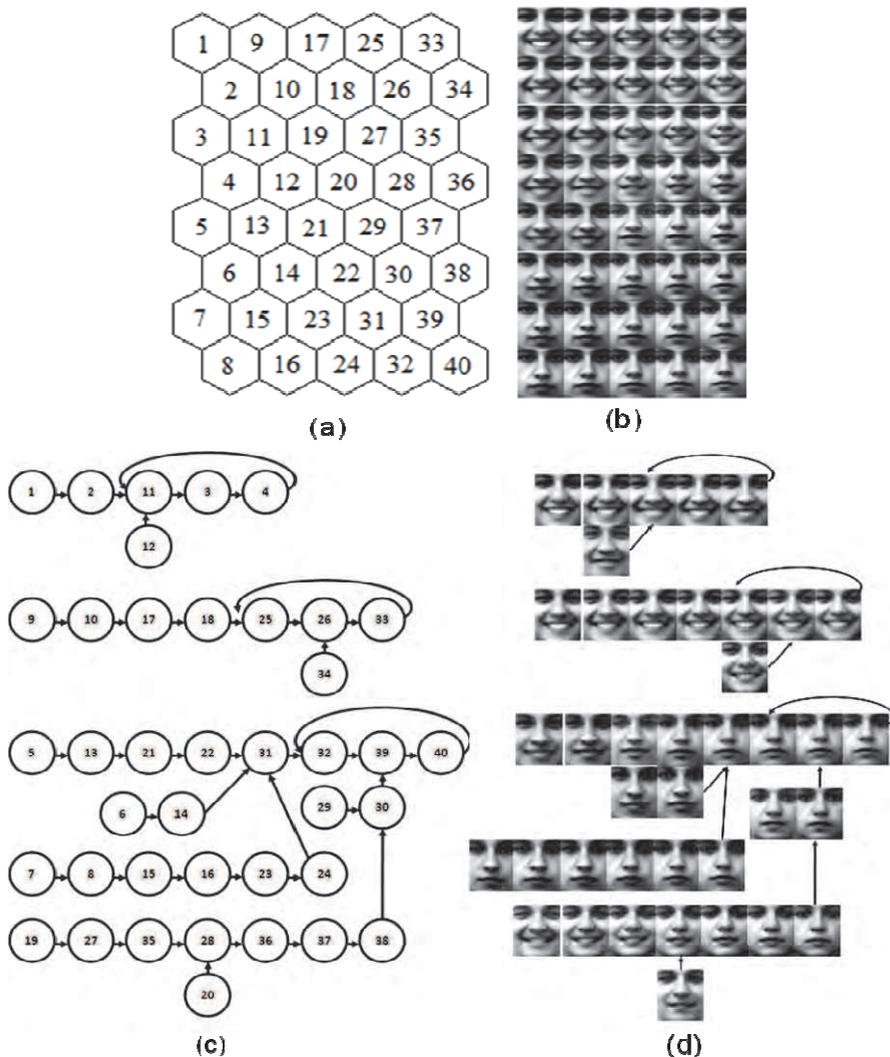


Fig. 6. Exploring the similarities and discontinuities of the high dimensional image face space composed of the whole set of frontal face images of the FEI database: standard SOM of size 8x5 (top left); visualization of the SOM neurons (top right); SOMM algorithm navigation (bottom left); visualization of the SOMM clustering (bottom right).

differences related to gender and ethnicity features. However, not all these pathways are feasible due to the discontinuities of the high dimensional and sparse image face space. In fact, as described by the SOMM algorithm, there are only five clusters possible to move along based on the principle of the locally optimal path. Therefore, although the standard SOM can explain the general information extracted by its neurons, its intrinsic self-organized manifolds have been only explicitly explained by the SOMM algorithm.

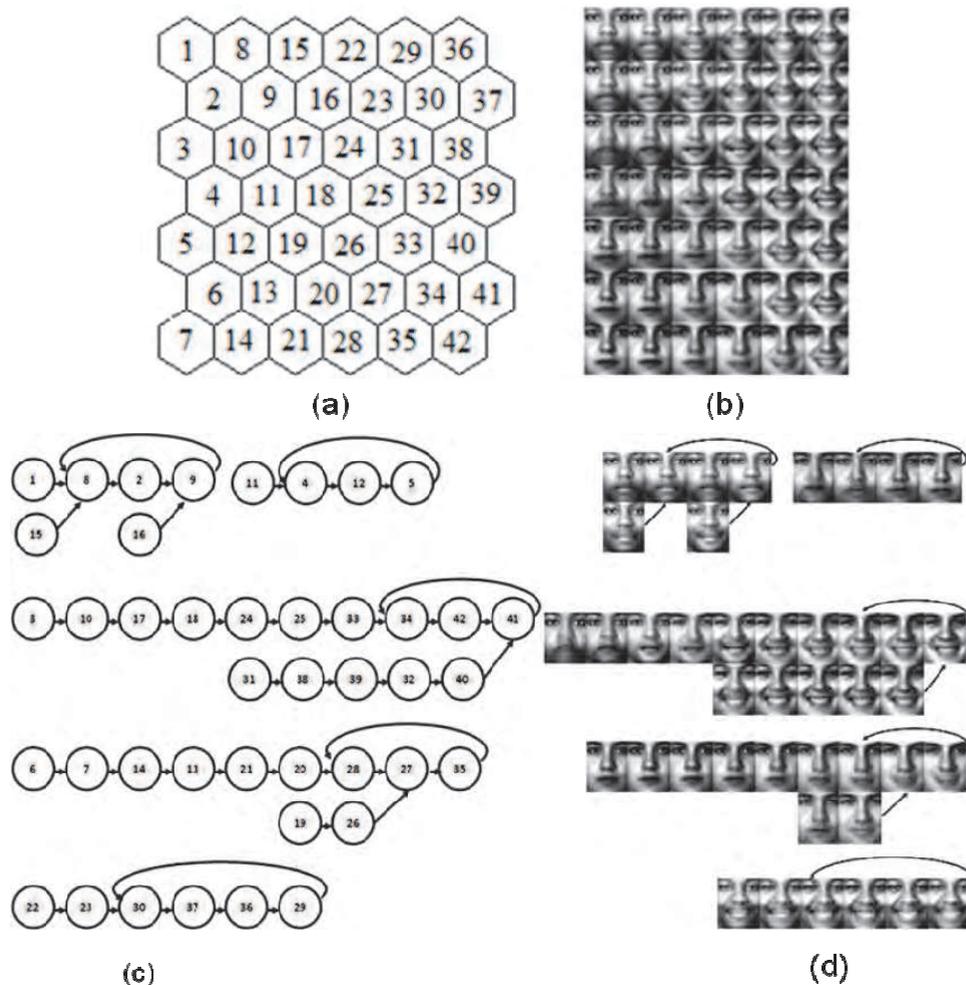


Fig. 7. Exploring the similarities and discontinuities of the high dimensional image face space composed of some frontal face images of the FERET database: standard SOM of size 7x6 (top left); visualization of the SOM neurons (top right); SOMM algorithm navigation (bottom left); visualization of the SOMM clustering (bottom right).

6. Conclusion

In this chapter, we proposed and implemented a self-organized manifold mapping algorithm that allows a better understanding of the information captured by the standard SOM neurons. The method is able not only to identify and explain the nature of the clusters defined by the SOM manifolds, but also to represent all the SOM neurons and their corresponding similarities and dissimilarities on the original data space. To describe the possible self-organized pathways to navigate on the high dimensional and sparse image face

space, we constructed a neighbourhood graph on the SOM neurons based on the principle of the locally optimal path. Such graph visualization method explicitly provides information about the number of clusters that describes the sample data under investigation, as well as the specific features extracted and explained by them. We believe that the algorithm proposed might be a powerful tool in SOM analysis, providing an intuitive explanation of the topologically constrained manifolds modelled by SOM and highlighting some perceptual properties commonly present in well-framed face image analysis such as facial expression, ethnicity and gender.

7. Acknowledgment

Portions of research in this paper use subsets of the FERET database of facial images collected under the FERET program.

8. References

- Anderson, J. R. (2005). *Cognitive Psychology and Its Implications*, 6th edition, Worth Publishers, New York USA.
- Asby, W. R. (1962). In *principles of self organization*, H.von Forester & G.W. Zopf eds., pp 255-278, London UK.
- Atlan, H. (1974). On a formal definition of organization, *Journal of theory in biology*, vol.45, pp. 295-304.
- Bakker, A., Kirwan, C. B., Miller, M., Stark, C. E. L. (2008). Pattern separation in the human Hippocampal CA3 and Dentate Gyrus, *Science Magazine*, vol. 319, pp. 1640.
- Bauer, H. U., Pawelzik, K. R. (1992). Quantifying the neighborhood preservation of Self-Organizing Feature Maps, *IEEE Transaction on Neural Networks*, vol. 3, no. 4, pp 570-579.
- Bear, M. F., Connors, B. W., Paradiso, M. A., (2007). *Neuroscience. Exploring the brain*. Lippincott Williams & Wilkins, 3rd ed.
- Brady, T. F., Konkle, T., Alvarez, G. A., Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details, *PNAS, Proceedings of the National Academy of Science of the United States of America*, vol. 105, no. 38, pp. 14325-14329.
- Brugger, D., Bogdan, M., Rosentiel, W. (2008) Automatic cluster detection in Kohonen's SOM, *IEEE Transaction on Neural Networks*, vol. 19, no. 3, pp. 442-459.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C. (2001). *Introduction to Algorithms*, 2nd ed. MIT Press.
- Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*, 2nd ed. Academic Press.
- Giraldi, G. A., Rodrigues, P. S., Kitani, E. C., Sato, J. R., Thomaz, C. E. (2008). Statistical Learning Approaches for Discriminant Features Selection, *Journal of the Brazilian Computer Society*, 14(2), 7-22.
- Kandel, E. R., Schwartz, J. H., Jessell, T. M. (2000). *Principles of Neural Science*, 4th ed., McGraw-Hill.
- Kitani, E. C., Hernandez, E. Del. Moral, Thomaz, C. E., Silva, L. A. (2010). Visual Interpretation of Self-Organizing Maps, *Proceedings of Neural Networks Brazilian (SBRN)*, IEEE CS Press, 37-42.

- Kitani, E. C., Thomaz, C. E., Gillies, D. F. (2006). A Statistical Discriminant Model for Face Interpretation and Reconstruction, Proceedings of Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI), IEEE CS Press, 247-254.
- Kiviluoto, K. (1995). Topology preservation in self-organizing maps, IEEE International Conference on Neural Networks, vol. 1, pp. 294-299, Washington, DC.
- Kohonen, T. (1982). Self-Organization and Associative Memory, Springer Verlag, Berlin.
- Kohonen, T. (1990). The Self-Organizing Map Proceedings of the IEEE, vol. 78, no.9.
- Kuhl, B. A., Shah, A., DuBrow, S., Wagner, A. D. (2010). Resistance to forgetting associated with hippocampus mediated reactivation during new learning, Nature Neuroscience, v. 13, no. 4.
- Mayer, R., Rauber, A. (2010). Visualizing clusters in Self-Organizing with minimum Spanning Trees, ICANN'10, Proceedings of the 20th International Conference on Artificial Neural Networks, pp. 426-431.
- Norman, K. A., O'Reilly, R. C. (2003). Modeling Hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach, Journal of Psychological Review, vol. 110, pp. 611-646.
- O'Reilly, R. C., McClelland J. L., (1994). Hippocampal conjunctive encoding, storage and recall: avoiding a trade off, Hippocampus vol. 4, pp. 661-682.
- O'Reilly, R. C., Rudy, J. W. (2001). Conjunctive representation in learning and memory: Principles of cortical and Hippocampal function, Journal of Psychological Review, vol. 108, pp. 311-345.
- Oja, E., (1982). A simplified neuron model as a Principal Component Analyser, Journal of Mathematical Biology, vol. 15, pp. 267-273.
- Philips, P. J., Wechsler, H., Huang, J., Rauss, P. (1998). The FERET database and evaluation procedure for face recognition algorithms, Image and Vision Computing, 16(5), 295-306.
- Pözlbauer, G., Rauber, A., Dittenbach, M. (2005). Graph projection techniques for Self Organizing Maps, ESANN'2005, European Symposium on Artificial Neural Networks, pp. 533-538.
- Purves, D., Augustine, G., Fitzpatrick, D., Katz, L. C., LaMantia, A. S., McNamara, J. O., Willians, S. M. (2001), Neuroscience, 2nd ed., Sinauer Associates.
- Rolls, E. T., (2007). An attractor network in the hippocampus: Theory and neurophysiology, Learning and Memory, vol. 14, pp. 714-731.
- Thomaz, C. E., Giraldo, G. A. (2010). A new ranking method for Principal Components Analysis and its application to face image analysis, Image and Vision Computing, vol. 28, no. 6, pp. 902-913.
- Thomaz, C. E., Amaral, V., Giraldo, G. A., Kitani, E. C., Sato, J. R., Gillies, D. F. (2009). A multi-linear discriminant analysis of 2D frontal face images, Proceedings of Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI), IEEE CS Press, 216-223.
- Treves, A., Rolls, E. T. (1994). Computational analysis of the role of the Hippocampus in memory, Hippocampus, vol. 4, pp. 374-391.
- Utsch, A. (2003). Maps for visualization of high-dimensional data space, in Proc. of Workshop on Self Organizing Maps, WSOM03, pp.225-230.
- Vesanto, J. (1999). Self-organizing map in Matlab: the SOM Toolbox, Proc. of the Matlab DSP Conference, Finland, pp. 35-40.



New Approaches to Characterization and Recognition of Faces

Edited by Dr. Peter Corcoran

ISBN 978-953-307-515-0

Hard cover, 252 pages

Publisher InTech

Published online 01, August, 2011

Published in print edition August, 2011

As a baby, one of our earliest stimuli is that of human faces. We rapidly learn to identify, characterize and eventually distinguish those who are near and dear to us. We accept face recognition later as an everyday ability. We realize the complexity of the underlying problem only when we attempt to duplicate this skill in a computer vision system. This book is arranged around a number of clustered themes covering different aspects of face recognition. The first section presents an architecture for face recognition based on Hidden Markov Models; it is followed by an article on coding methods. The next section is devoted to 3D methods of face recognition and is followed by a section covering various aspects and techniques in video. Next short section is devoted to the characterization and detection of features in faces. Finally, you can find an article on the human perception of faces and how different neurological or psychological disorders can affect this.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Edson C. Kitani, Emilio M. Hernandez, Gilson A. Giraldo and Carlos E. Thomaz (2011). Exploring and Understanding the High Dimensional and Sparse Image Face Space: a Self-Organized Manifold Mapping, *New Approaches to Characterization and Recognition of Faces*, Dr. Peter Corcoran (Ed.), ISBN: 978-953-307-515-0, InTech, Available from: <http://www.intechopen.com/books/new-approaches-to-characterization-and-recognition-of-faces/exploring-and-understanding-the-high-dimensional-and-sparse-image-face-space-a-self-organized-manif>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.