

TopoART: A Topology Learning Hierarchical ART Network

Marko Tscherepanow

Bielefeld University, Applied Informatics
Universitätsstraße 25, 33615 Bielefeld, Germany
marko@techfak.uni-bielefeld.de

Abstract. In this paper, a novel unsupervised neural network combining elements from Adaptive Resonance Theory and topology learning neural networks, in particular the Self-Organising Incremental Neural Network, is introduced. It enables stable on-line clustering of stationary and non-stationary input data. In addition, two representations reflecting different levels of detail are learnt simultaneously. Furthermore, the network is designed in such a way that its sensitivity to noise is diminished, which renders it suitable for the application to real-world problems.

Keywords: on-line learning, topology learning, Adaptive Resonance Theory, hierarchical clustering

1 Introduction

For numerous tasks, the traditional off-line learning approach with separate training, validation, and test phases is not sufficient. The diagnosis of genetic abnormalities [1], interactive teaching of a humanoid robot [2], and the subcellular localisation of proteins [3] constitute some examples for such problems. As a consequence, incremental on-line learning has become more popular in recent years, since such machine learning techniques are required to gradually complete knowledge or adapt to non-stationary input distributions.

In this paper, a novel neural network is introduced which combines incremental and fast on-line clustering with topology learning. Based on its origins in Adaptive Resonance Theory (ART) networks, it is called TopoART. Thanks to these origins, TopoART creates stable representations while retaining its ability to learn new data. In order to render TopoART more suitable for real-world applications, it was designed in such a way that it becomes insensitive to noise. Furthermore, TopoART creates a hierarchical representation of the input distribution reflecting different levels of detail.

In Sect. 2, related learning methods are outlined. Afterwards, details of the TopoART algorithm are introduced in Sect. 3. Then, clustering results for artificial and real-world data are presented in Sect. 4. Here, the ability of TopoART to cope with noise and to incrementally learn new input data from non-stationary distributions will be demonstrated as well. Finally, Sect. 5 summarises the most important points of this paper.

2 Related Work

The k-means algorithm [4], which constitutes a very well-known unsupervised learning technique, determines a partitioning of an input distribution into k regions or rather clusters. Each cluster is represented by a reference vector. The determination of the number of required clusters constitutes a crucial problem. For this reason, the Linde-Buzo-Gray (LBG) algorithm [5] was developed. Based on a fixed training set, it successively computes sets of reference vectors of increasing size until a stopping criterion is fulfilled. The topological structure of the input data is not considered by this algorithm.

In 1982, the Self-Organising Feature Maps (SOFMs) [6], which map input data to a lattice of neurons, were introduced. Here, the reference vectors are encoded by the weights of the neurons. The lattice possesses a predefined topological structure, the dimension of which is usually lower or equal to the dimension of the input space. If the input distribution is not known in advance, an appropriate lattice structure is very difficult to choose. This problem was solved by the Growing Neural Gas (GNG) algorithm [7]. It allows for the incremental incorporation of new neurons and the learning of the input distribution's topology by adding and deleting edges between different neurons. But the representation of the input distribution is not stable: A continuing presentation of input data, even if they have already been learnt, results in a continuing adaptation of the neurons' weights, i.e. the reference vectors, and the network topology. As a consequence, already acquired knowledge gets lost due to further training. This problem has been termed the *stability-plasticity dilemma* [8]. It occurs, for instance, if the input distribution is complex. But small changes of the input probabilities or the sequencing of the input data may cause a similar effect.

Adaptive Resonance Theory (ART) networks have been proposed as a solution to the stability-plasticity dilemma [8]. These networks learn top-down expectations which are matched with bottom-up input. The expectations, which are called categories, summarise sets of input data to clusters. Depending on the type of ART network, the categories exhibit different shapes such as a hyperspherical shape [9], an elliptical shape [10], or a rectangular shape [11]. Besides enabling ART networks to create stable and plastic representations, the categories allow for an easy novelty detection. But in contrast to SOFMs and GNG, ART networks do not capture the topology of the input data. Furthermore, their ability of stable learning leads to an increased sensitivity to noise.

In 2006, the Self-Organising Incremental Neural Network (SOINN) was introduced [12]. Similar to GNG, SOINN clusters input data by incrementally adding neurons, the weights of which represent reference vectors, and representing the topology by edges between nodes. But it has several additional features: Firstly, SOINN has a two-layered structure representing the input distribution at different levels of detail. Additionally, this structure decreases the sensitivity to noise. The second layer is trained after the training of the first layer has been finished. Secondly, novelty detection can be performed based on an adaptive threshold. Thirdly, each neuron has an individual learning rate which decays if the amount of input samples it represents increases. By this means, a more stable represen-

tation is achieved. But the weights of the neurons do not stabilise completely. Furthermore, a high number of relevant parameters (8 parameters per layer) has to be selected in order to apply SOINN.

The Enhanced Self-Organising Incremental Neural Network (ESOINN) [13] solves some of the above mentioned problems: By removing the second layer and one condition for the insertion of new neurons, the number of required parameters is considerably decreased (4 in total). Furthermore, the whole network can be trained on-line. But similar to SOINN, the weights do not stabilise completely. Moreover, ESOINN loses the ability to create hierarchical representations.

TopoART combines the advantages of ART and topology learning networks. From its ART ancestors, it inherits the ability of fast and stable on-line learning using expectations (categories). But the categories are extended by edges reflecting the topology of the input distribution enabling the formation of arbitrarily shaped clusters. In addition, it adopts the ability to represent input data at different levels of detail from SOINN; but unlike SOINN, TopoART learns both levels simultaneously.

3 The TopoART Algorithm

The basic structure and the computational framework of TopoART are strongly related to Fuzzy Art [11] – a very efficient ART network utilising rectangular categories. Fuzzy ART constitutes a three-layered neural network (see Fig. 1). Input is presented to the initial layer $F0$. Here, the input vectors $\underline{x}(t)$ are encoded using complement coding. The encoded input vectors are denoted by $\underline{x}^{F1}(t)$. As a consequence of the usage of complement coding, each component $x_i(t)$ of an input vector $\underline{x}(t)$ has to lie in the interval $[0, 1]$.

$$\underline{x}(t) = [x_1(t), \dots, x_d(t)]^T \quad (1)$$

$$\underline{x}^{F1}(t) = [x_1(t), \dots, x_d(t), 1 - x_1(t), \dots, 1 - x_d(t)]^T \quad (2)$$

The encoded input vectors $\underline{x}^{F1}(t)$ are transmitted to the comparison layer $F1$. From here, they activate the neurons of the representation layer $F2$:

$$z_i^{F2}(t) = \frac{|\underline{x}^{F1}(t) \wedge \underline{w}_i^{F2}(t)|_1}{\alpha + |\underline{w}_i^{F2}(t)|_1} \quad (3)$$

The activation $z_i^{F2}(t)$ (choice function) constitutes a measure for the similarity between $\underline{x}^{F1}(t)$ and the category represented by neuron i . $|\cdot|_1$ and \wedge denote the city block norm and a component-wise minimum operation, respectively. The parameter α must be set slightly higher than zero. The choice of the actual value is not crucial.¹ In general, $z_i^{F2}(t)$ prefers small categories to large ones.

After all $F2$ neurons have been activated, the best-matching neuron bm , i.e. the neuron with the highest activation, is selected. But the category represented

¹ α was set to 0.001 for all experiments presented in this paper.

by its weights $\underline{w}_{bm}^{F2}(t)$ is only allowed to grow and enclose a presented input vector if resonance occurs, i.e. if the match function (4) is fulfilled.

$$\frac{|\underline{x}^{F1}(t) \wedge \underline{w}_{bm}^{F2}(t)|_1}{|\underline{x}^{F1}(t)|_1} \geq \rho \quad (4)$$

In case of resonance, the weights \underline{w}_{bm}^{F2} of the chosen neuron are adapted and the output $\underline{y}(t)$ of the network is set:

$$\underline{w}_{bm}^{F2}(t+1) = \beta(\underline{x}^{F1}(t) \wedge \underline{w}_{bm}^{F2}(t)) + (1-\beta)\underline{w}_{bm}^{F2}(t) \quad (5)$$

$$y_i(t) = \begin{cases} 0 & \text{if } i \neq bm \\ 1 & \text{if } i = bm \end{cases} \quad (6)$$

β denotes the learning rate. Setting β to 1, trains the network in fast-learning mode; i.e., each learnt input is enclosed by the category that matches it best. As the categories cannot shrink according to (5), the formed representations are stable. The current size $S_i(t)$ of a category can be derived from the weights $\underline{w}_i^{F2}(t)$ of the corresponding neuron i :

$$S_i(t) = \sum_{j=1}^d \left| (1 - w_{i,d+j}^{F2}(t)) - w_{i,j}^{F2}(t) \right| \quad (7)$$

The growth of the categories is limited by the vigilance parameter ρ and the dimension of the input space d , which determine the maximum category size S^{\max} .

$$S^{\max} = d(1 - \rho) \quad (8)$$

Assuming a neuron was not able to fulfil (4), its activation is reset. Then a new best-matching node is chosen. If no suitable best-matching neuron is found, a new neuron representing $\underline{x}(t)$ is incorporated and resonance occurs.

TopoART is composed of two Fuzzy ART-like components (TopoART a and TopoART b) using a common $F0$ layer performing complement coding (see Fig. 1). But in contrast to Fuzzy ART, each $F2$ neuron i of both components has a counter denoted by n_i^a and n_i^b , respectively, which counts the number of input samples it has learnt. An encoded input vector is only propagated to TopoART b if resonance of TopoART a occurred and $n_{bm}^a \geq \phi$. Every τ learning cycles, all neurons with a counter smaller than ϕ are removed. Therefore, such neurons are called node candidates. Once n_i equals or surpasses ϕ , the corresponding neuron cannot be removed anymore; i.e., it becomes a permanent node. By means of this procedure, the network becomes insensitive to noise but is still able to learn stable representations.

In order to allow for fast on-line learning, the learning rate β_{bm} used for adapting the weights of the best-matching neuron is always set to 1 yielding:

$$\underline{w}_{bm}^{F2}(t+1) = \underline{x}^{F1}(t) \wedge \underline{w}_{bm}^{F2}(t) \quad (9)$$

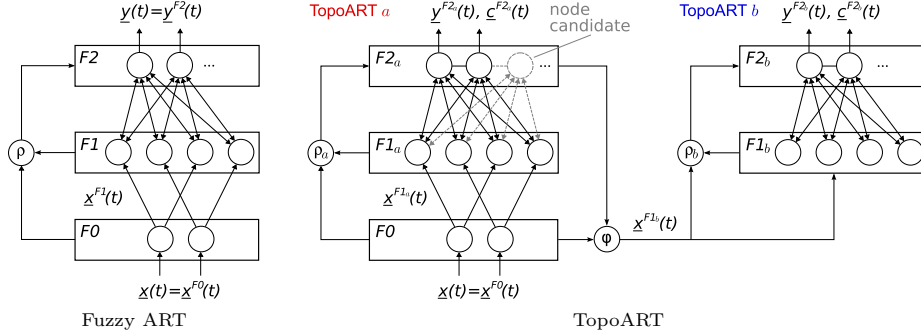


Fig. 1. Comparison of Fuzzy ART and TopoART. TopoART consists of two Fuzzy ART-like components using the same input layer. Furthermore, the $F2$ nodes of TopoART are connected by edges defining a topological structure. In order to reduce its sensitivity to noise, TopoART evaluates the benefit of neurons (node candidates) before they are fully incorporated.

Rather than only determining the best-matching neuron bm and modifying its weights, the neuron sbm with the second highest activation that fulfils (4) is adapted as well (10). Here, β_{sbm} should be chosen smaller than 1, as neuron sbm – in contrast to neuron bm – is only intended to partly learn $\underline{x}^{F1}(t)$.

$$\underline{w}_{sbm}^{F2}(t+1) = \beta_{sbm}(\underline{x}^{F1}(t) \wedge \underline{w}_{sbm}^{F2}(t)) + (1 - \beta_{sbm})\underline{w}_{sbm}^{F2}(t) \quad (10)$$

As a result of this procedure, the insensitivity to noise is further increased, since the categories are more likely to grow in relevant areas of the input space.

If $\phi=1$ and $\beta_{sbm}=0$, inserted nodes immediately become permanent, all input vectors are propagated to TopoART b , and only the weights of the best-matching neuron bm are adapted during a learning cycle. In this case, the categories of TopoART a and TopoART b equal the categories of Fuzzy ART networks trained in fast-learning mode with the vigilance parameters ρ_a and ρ_b , respectively.

In order to enable TopoART to learn topologies, a lateral connection or rather edge between the neurons bm and sbm is created, if a second best-matching neuron can be found. These edges define a topological structure. They are not used for activating other neurons. If the neurons bm and sbm have already been connected by an edge, it remains unchanged, since the edges do not possess an age parameter in contrast to the edges in ESOINN, SOINN, and GNG. They are only removed, if one of the adjacent neurons is removed. As a consequence, edges between permanent nodes are stable. But new edges can still be created. This mechanism constitutes an extension of Fuzzy ART's solution to the stability-plasticity dilemma, which enables the representation of new input while retaining the already learnt representations.

In order to refine the representation of TopoART a by means of TopoART b , ρ_b should be higher than ρ_a . Within the scope of this paper, ρ_b is determined according to (11), which diminishes the maximum category size S^{\max} by 50%.

$$\rho_b = \frac{1}{2}(\rho_a + 1) \quad (11)$$

In this way, TopoART b learns a more detailed representation which is less influenced by noise. Connections between categories of TopoART a can be split in TopoART b resulting in a hierarchical representation of the input data.

The output $\underline{y}(t)$ of both TopoART components can be computed in a similar way as with Fuzzy ART. In addition, the connected components or rather clusters are labelled (cf. [12] and [13]): Starting from an unlabelled neuron, all connected neurons receive a specific label. Then, a new unlabelled neuron is searched for. This process is repeated until no unlabelled neurons remain. The vector $\underline{c}(t)$ provides the cluster labels of all $F2$ neurons. For reasons of stability, only permanent nodes are considered for the computation of $\underline{y}(t)$ and $\underline{c}(t)$. By this, the clusters can grow and fuse, but they are prevented from shrinking.

4 Results

TopoART was evaluated using three different types of data: (i) stationary artificial data, (ii) non-stationary artificial data, and (iii) stationary real-world data. As stationary artificial input distribution, a dataset copying the one used for the evaluation of SOINN [12] was chosen (see Fig. 2(a)). It comprises two Gaussian components (A and B), two ring-shaped components (C and D), and a sinusoidal component (E) composed from three subcomponents (E1, E2, and E3). Each component encompasses 18,000 individual samples. Additionally, the input distribution includes uniformly distributed random noise amounting to 10% of the total sample number (10,000 samples).

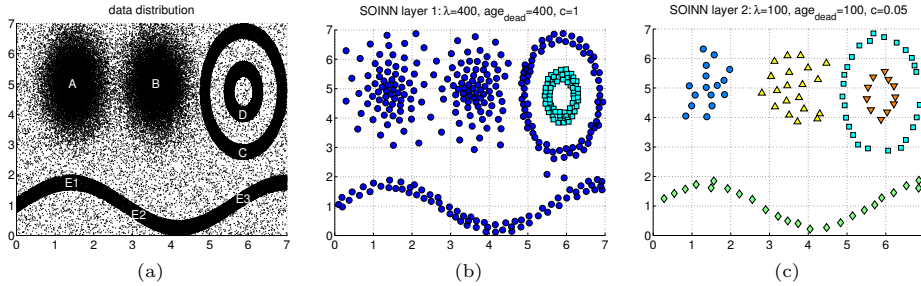


Fig. 2. Input distribution and clustering results of SOINN. The input distribution (a) is successfully clustered by SOINN. Here, the representation is refined from layer 1 (b) to layer 2 (c). Reference vectors belonging to the same cluster share a common colour and symbol.

This dataset was used to train a SOINN system. The results of both of its layers are depicted in Fig. 2(b) and Fig. 2(c), respectively. The values for λ ,

age_{dead} , and c were selected in such a way that results comparable to those published in [12] were achieved. In contrast, the settings for α_1 , α_2 , α_3 , β , and γ were directly adopted from [12] for both layers $(\frac{1}{6}, \frac{1}{4}, \frac{1}{4}, \frac{2}{3}, \frac{3}{4})$.

Figure 2 shows that SOINN was able to create a hierarchical representation of the input distribution: The two clusters in layer 1 were refined in layer 2 which distinguishes five clusters. Additionally, the second layer reduced the influence of noise.

The dataset shown in Fig. 2(a) cannot be used directly to train a Fuzzy ART or a TopoART network. Rather, all samples have to be scaled to the interval $[0, 1]$ so as to render them compatible with complement coding.

As Fuzzy ART constitutes the basis of TopoART, it was analysed first. For comparison reasons, β was set to 1 like β_{bm} in TopoART. ρ was selected in such a way as to roughly fit the thickness of the elliptic and sinusoidal components of the input distribution. As this network does not possess any means to decrease the sensitivity to noise, virtually the whole input space was covered by categories (see Fig. 3(a)).

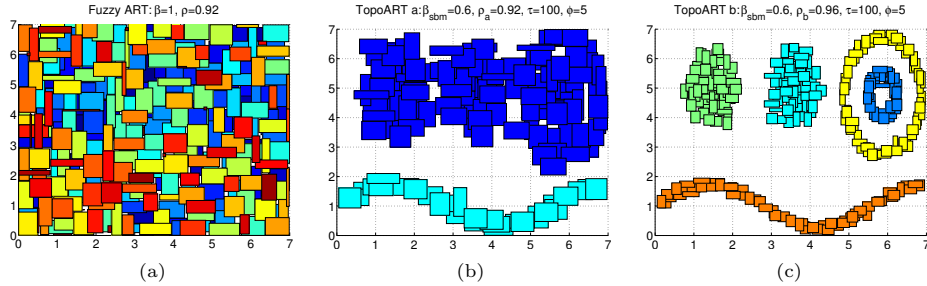


Fig. 3. Comparison of the ART networks' results. Fuzzy ART (a) learnt rectangular categories covering virtually the complete input space. In contrast, TopoART learnt a noise insensitive representation in which the categories have been summarised to arbitrarily shaped clusters. The representation of TopoART *a* (b) was further refined by TopoART *b* (c). Here, all categories of an individual cluster are painted with the same colour.

In contrast to Fuzzy ART, both TopoART components created representations reflecting the relevant regions of the input distribution very well (see Figs. 3(b) and 3(c)). This is remarkable since the value of ρ_a was equal to the value of the vigilance parameter ρ of the Fuzzy ART network. Similar to the layers of SOINN, the representation of TopoART was refined from TopoART *a* to TopoART *b*: While TopoART *a* comprises two clusters, TopoART *b* distinguishes five clusters corresponding to the five components of the input distribution. By virtue of the filtering of samples by TopoART *a* and due to the fact that ρ_b is higher than ρ_a (cf. (11)), the categories of TopoART *b* reflect the input distribution in more detail. This property is particularly useful, if small areas of the input space have to be clustered with high accuracy. Here, TopoART *a*

could filter input from other regions and TopoART b could create the desired detailed representation.

The learning of non-stationary data was also investigated using the input distribution shown in Fig. 2(a). But now, input from different regions was presented successively (see Fig. 4).

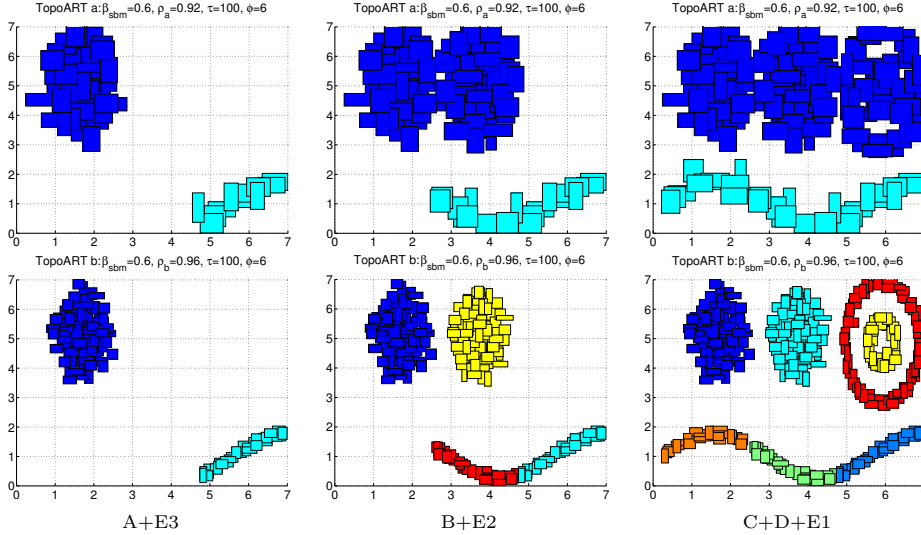


Fig. 4. Training of TopoART with non-stationary artificial data. A TopoART network was consecutively trained using three different input distributions: A+E3, B+E2, and C+D+E1. Each column shows the results after finishing training with data from the given regions. All categories of an individual cluster are painted with the same colour.

Figure 4 shows that both components of TopoART incrementally learnt the presented input. Already created representations remained stable when the input distribution changed. Similar to the stationary case, TopoART b performed a refinement of the representation of TopoART a . But here, the sub-regions E1, E2, and E3 were separated, since the corresponding input samples were presented independently and could not be linked. TopoART a was able to compensate for this effect, as its lower vigilance parameter ρ_a allowed for larger categories which could form connections between the sub-regions.

Finally, a dataset originally used to investigate methods for the direct imitation of human facial expressions by the user-interface robot iCat was applied [14]. From this dataset, the images of all 32 subjects (12 female, 20 male) that were associated with predefined facial expressions were selected, which resulted in a total of 1783 images. These images had been acquired using two lighting conditions: daylight and artificial light. They were cropped, scaled to a size of 64×64 pixels, and successively processed by principal component analysis keeping 90% of the total variance (cf. Fig. 5(a)).

After training a TopoART network with these data, the resulting clusters were compared to the partitionings based on labels reflecting the subjects and the lighting conditions. Here, two standard measures were used: the Jaccard coefficient J and the Rand index R [15]. Both provide values between 0 and 1, with higher values indicating a higher degree of similarity. As always more than one criterion (subjects, lighting conditions, facial expressions, gender, usage of glasses, etc.) for partitioning these images influences the clustering process, values considerably lower than 1 may also indicate similarity. The splitting of clusters caused by these additional criteria, for example, results in a decrease of the Jaccard coefficient.

Due to the filtering process controlled by ϕ , several input samples were regarded as noise and excluded from cluster formation. Within the scope of the evaluation, they were associated with the cluster containing the permanent node with the highest activation. Since the original activation (3) depends on the category size, an alternative activation (12) was used to analyse the formed clusters.² It constitutes a normalised version of the city block distance between an input sample and the respective category, which was successfully applied in [3] and [16].

$$z_i^{F2}(t) = 1 - \frac{|(\underline{x}^{F1}(t) \wedge \underline{w}_i^{F2}(t)) - \underline{w}_i^{F2}(t)|_1}{d} \quad (12)$$

Figure 5 shows the results for both TopoART components. Here, the parameter τ was adopted from the previous experiments; β_{sbm} , ϕ , and ρ_a were iterated over representative values from the relevant intervals and selected in such a manner as to maximise the Jaccard coefficient for the partitioning according to the subjects. Then β_{sbm} and ϕ were fixed, while the influence of ρ_a is illustrated. Here, no results for $\rho_a < 0.75$ are displayed, since a further decrease of ρ_a does not cause any changes.

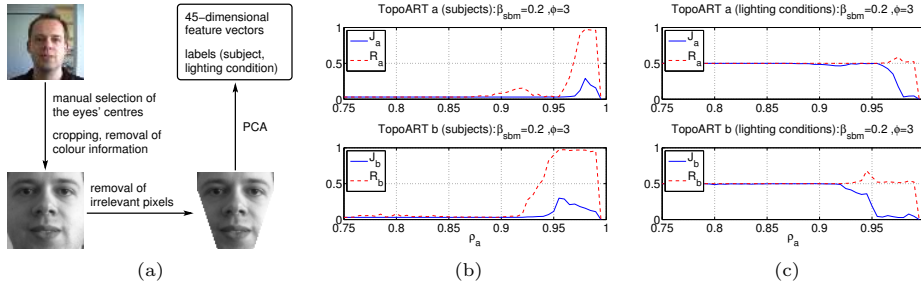


Fig. 5. Clustering of stationary real-world data. After preprocessing facial images (a), they were clustered by a TopoART network. Then, it was analysed how accurately the resulting clusters reflect the different subjects (b) and the two lighting conditions (c).

² The original activation (3) was still utilised for training.

According to Fig. 5(b), the partitioning with respect to the 32 subjects is best represented if ρ_a is high. Here, J_b and R_b increase for lower values of ρ_a than J_a and R_a , which reflects the finer clustering of TopoART b . In contrast, the partitioning according to the two lighting conditions (cf. Fig. 5(c)) is impaired for high values of ρ_a , which is indicated by the decrease of J_a and J_b . The more detailed representation of TopoART b amplifies this effect. Choosing ρ_a appropriately, e.g. $\rho_a=0.96$, the representations of both partitionings can be combined by a single TopoART network. In this case, TopoART a represents the coarser partitioning with respect to the lighting conditions and TopoART b the finer partitioning according to the subjects.

For the results given in Fig. 5, the relevant parameters were selected in such a way as to maximise the Jaccard coefficient for the partitioning according to the subjects. Hence, the representation of the lighting conditions is not optimal. If the parameters are optimised to provide the maximum Jaccard coefficient for this partitioning, the following maximum values are reached: $J_a=0.779$ and $J_b=0.671$. This confirms the previous results according to which TopoART a creates a less detailed representation of the input data which is more suited to represent the two lighting conditions.

5 Conclusion

TopoART – the neural network proposed in this paper – successfully combines properties from ART and topology learning approaches: The categories originating from ART systems are connected by means of edges. In this way, clusters of arbitrary shapes can be formed. In addition, a filtering mechanism decreases the sensitivity to noise. Similar to SOINN, representations exhibiting different levels of detail are formed. But TopoART enables parallel learning at both levels requiring only 4 parameters ($\beta_{sbm}, \phi, \rho_a, \tau$) to be set, which constitutes a reduction of 75% compared to SOINN. Moreover, representations created by TopoART are stable.

Although the number of parameters to be set is small, they must be chosen using some kind of prior knowledge about the input data distribution. This task is very difficult, in particular, as TopoART is trained on-line. In order to solve this problem, the hierarchical structure of TopoART can be exploited, since it provides alternative clusterings of the input data distribution. By means of interaction during the learning process, these clusterings could be evaluated with respect to the current task or other criteria.

The capability of TopoART to capture hierarchical relations and the topology of presented data might be of interest for numerous tasks, e.g. the representation of complex sensory and semantic information in robots. In principle, TopoART could even be extended to a multi-level structure that captures hierarchical relations more comprehensively.

Acknowledgements. This work has partly been supported by the German Research Foundation (DFG) within the Center of Excellence ‘Cognitive Interaction

Technology’, Research Area D (Memory and Learning). The author particularly wishes to thank Sina Kühnel and Marc Kammer for many fruitful discussions during the preparation of this manuscript.

References

1. Vigdor, B., Lerner, B.: Accurate and fast off and online fuzzy ARTMAP-based image classification with application to genetic abnormality diagnosis. *IEEE Transactions on Neural Networks* **17**(5) (2006) 1288–1300
2. Goerick, C., Schmüdderich, J., Bolder, B., Janßen, H., Gienger, M., Bendig, A., Heckmann, M., Rodemann, T., Brandl, H., Domont, X., Mikhailova, I.: Interactive online multimodal association for internal concept building in humanoids. In: *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*. (2009) 411–418
3. Tscherepanow, M., Jensen, N., Kummert, F.: An incremental approach to automated protein localisation. *BMC Bioinformatics* **9**(445) (2008)
4. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*. Volume 1. (1967) 281–297
5. Linde, Y., Buzo, A., Gray, R.M.: An algorithm for vector quantizer design. *IEEE Transactions on Communications* **COM-28** (1980) 84–95
6. Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biological Cybernetics* **43**(1) (1982) 59–69
7. Fritzke, B.: A growing neural gas network learns topologies. In: *Neural Information Processing Systems*. (1994) 625–632
8. Grossberg, S.: Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science* **11** (1987) 23–63
9. Anagnostopoulos, G.C., Georgiopoulos, M.: Hypersphere ART and ARTMAP for unsupervised and supervised incremental learning. In: *Proceedings of the International Joint Conference on Neural Networks*. Volume 6. (2000) 59–64
10. Anagnostopoulos, G.C., Georgiopoulos, M.: Ellipsoid ART and ARTMAP for incremental clustering and classification. In: *Proceedings of the International Joint Conference on Neural Networks*. Volume 2. (2001) 1221–1226
11. Carpenter, G.A., Grossberg, S., Rosen, D.B.: Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks* **4** (1991) 759–771
12. Furao, S., Hasegawa, O.: An incremental network for on-line unsupervised classification and topology learning. *Neural Networks* **19** (2006) 90–106
13. Furao, S., Ogura, T., Hasegawa, O.: An enhanced self-organizing incremental neural network for online unsupervised learning. *Neural Networks* **20** (2007) 893–903
14. Tscherepanow, M., Hillebrand, M., Hegel, F., Wrede, B., Kummert, F.: Direct imitation of human facial expressions by a user-interface robot. In: *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*. (2009) 154–160
15. Xu, R., Wunsch II, D.C.: *Clustering*. Wiley–IEEE Press (2009)
16. Tscherepanow, M.: Image analysis methods for location proteomics. PhD thesis, Faculty of Technology, Bielefeld University (2008)