# Associating SOM Representations of Haptic Submodalities

Magnus Johnsson and Christian Balkenius

*Abstract*— **We have experimented with a bio-inspired self-organizing texture and hardness perception system which automatically learns to associate the representations of the two submodalities with each other. To this end we have developed a microphone based texture sensor and a hardness sensor that measures the compression of the material at a constant pressure. The system is based on a novel variant of the Self-Organizing Map (SOM), called Associative Self-Organizing Map (A-SOM). The A-SOM both develops a representation of its input space and learns to associate this with the activity in an external SOM or A-SOM. The system was trained and tested with multiple samples gained from the exploration of a set of 4 soft and 4 hard objects of different materials with varying textural properties. The system successfully found representations of the texture and hardness submodalities and also learned to associate these with each other.**

## I. INTRODUCTION

An efficient multimodal perceptual system should be able to associate different modalities and submodalities with each other. This provides an ability to activate the subsystem for a modality even when its sensory input is limited or nonexistent as long as there are activities in subsystems for other modalities, which the subsystem has learned to associate with certain patterns of activity, which usually comes together with the patterns of activity in the other subsystems. For example, in humans the sensory information gained when the texture of an object is felt in the pocket can invoke visual images of the object or a feeling for its hardness.

To study the association of different submodalities we used the two haptic submodalities texture and hardness, thus gaining experience with these submodalities in robotics as well. There have been some previous studies of texture and hardness in robotics. For example, Hosoda et al [6] have built an anthropomorphic fingertip with distributed receptors consisting of two silicon rubber layers of different hardness. The silicon rubber layers contain two different sensors, strain gauges and polyvinylidene fluoride films, which yield signals that enabled the discrimination of five different materials pushed and rubbed by the fingertip. Mayol-Cuevas et al [14] describes a system for tactile texture recognition, which employs a sensing pen with a microphone that is manually rubbed over the explored materials. The system uses a supervised Learning Vector Quantization (LVQ) classifier system to identify with 93% accuracy 18 common materials after signal processing with the Fast Fourier Transform (FFT).

Magnus Johnsson is with the Department of Computer Science and Lund University Cognitive Science, Lund University, Sweden Magnus.Johnsson@lucs.lu.se

Christian Balkenius is with Lund University Cognitive Science, Lund University, Sweden Christian.Balkenius@lucs.lu.se

Edwards et al [4] have used a vinyl record player with the needle replaced with an artificial finger with an embedded microphone to quantify textural features by using a group of manufactured discs with different textural patterns. Campos and Bajcsy [3] explored haptic Exploratory Procedures (EPs) based on human haptic EPs proposed by Lederman and Klatzky, among them an EP for hardness exploration in which the applied force is measured for a given displacement.

We have done some previous experimentation with texture/hardness perception [11]. In this experiments we test our hardness and texture sensors together with a self-organizing systems that develops monomodal as well as bimodal representations of texture and hardness. Our other previous research on haptic perception has resulted in the design and implementation of a number of versions of three different working haptic systems. The first system [7] was a system for haptic size perception. It used a simple three-fingered robot hand, the LUCS Haptic Hand I, with the thumb as the only movable part. The LUCS Haptic Hand I was equipped with 9 piezo electric tactile sensors. This system used self-organizing maps (SOMs) [12] and a neural network with leaky integrators and it successfully learned to categorize a test set of spheres and cubes according to size.

The second system [8] was a system for haptic shape perception and used a three-fingered 8 d.o.f. robot hand, the LUCS Haptic Hand II, equipped with a wrist for horizontal rotation and a mechanism for vertical re-positioning. This robot hand was equipped with 45 piezo electric tactile sensors. This system used active explorations of the objects by several grasps with the robot hand to gather tactile information, which together with the positioning commands to the actuators (thus a kind of pseudoproprioception) were cross-coded by, depending on the version, either tensor product (outer product) operations or a novel neural network, the Tensor Multiple Peak SOM (T-MPSOM) [8]. The cross-coded information was categorized by a SOM. The system successfully learned to discriminate between different shapes as well as between different objects within a shape category when tested with a set of spheres, blocks and cylinders.

The third system [9] was a bio-inspired self-organizing system for haptic shape and size perception based solely on proprioceptive data from a 12 d.o.f. anthropomorphic robot hand with proprioceptive sensors [10]. The system was trained with 10 different objects of different sizes from two different shape categories and tested with both the training set and a novel set with 6 previously unused objects. It was able to discriminate the shape as well as the size of the objects in both the original training set and the set of new objects.

This paper explores a bio-inspired self-organizing texture and hardness perception system, which automatically learns to associate the representations of the two submodalities with each other. The system is based on a novel variant of the SOM, called Associative Self-Organizing Map (A-SOM), and it employs a microphone based texture sensor and a hardness sensor that measures the compression of the material at a constant pressure. The system is bio-inspired in the sense that it employs a variation of the SOM to represent the two submodalities texture and hardness, and the SOM shares many features with brain maps [13]. It is also bio-inspired in the sense that different submodaliteter are integrated. That different submodaliteter are integrated in unimodal association areas in the human brain is a well established fact [15]. The texture sensor is also bio-inspired. Our system is based on the transduction of vibrations from a metal edge and which are transmitted to a microphone. In humans the mechanoreceptors respond to vibrations as well [5].

## II. A-SOM

The A-SOM (Fig. 1) can be considered as a Self-Organizing Map (SOM) [12] which learns to associate the activity of an external A-SOM or SOM with its own activity. It consists of an $I \times J$ grid of neurons with a fixed number of neurons and a fixed topology. Each neuron $n_{ij}$ is associated with two weight vectors $w^a_{ij} \in R^n$ and $w^b_{ij} \in R^m$ where $m$ equals the number of neurons in an external A-SOM or SOM. $w^a_{ij}$ is initialized randomly to numbers between $0$ and $1$, whereas all elements of $w^b_{ij}$ are initialized to $0$.

At time $t$ each neuron $n_{ij}$ receives two normalized input vectors $x^a(t) \in R^n$ and $x^b(t) \in R^m$.

The neuron $c$ associated with the weight vector $w^a_c(t)$ most similar to the input vector $x^a(t)$ is selected:

$$c = \text{arg max}_c\{||x^a(t)w^a_c(t)||\} \quad (1)$$

The activity in the neuron $n_{ij}$ is given by

$$y_{ij}(t) = \left[y^{input}_{ij}(t) + y^{extern}_{ij}(t)\right]/2 \quad (2)$$

where

$$y^{input}_{ij}(t) = G(\|n_{ij} - c\|) \quad (3)$$

and

$$y^{extern}_{ij}(t) = x^b(t)w^b_{ij}(t) \quad (4)$$

$G()$ is a Gaussian function with $G(0) = 1$, and $\| \cdot \|$ is the Euclidean distance between two neurons.

The weights $w^a_{ijk}$ are adapted by

$$w^a_{ijk}(t+1) = w^a_{ijk}(t) + \alpha(t)G_{ijc}(t)\left[x^a_k(t) - w^a_{ijk}(t)\right] \quad (5)$$

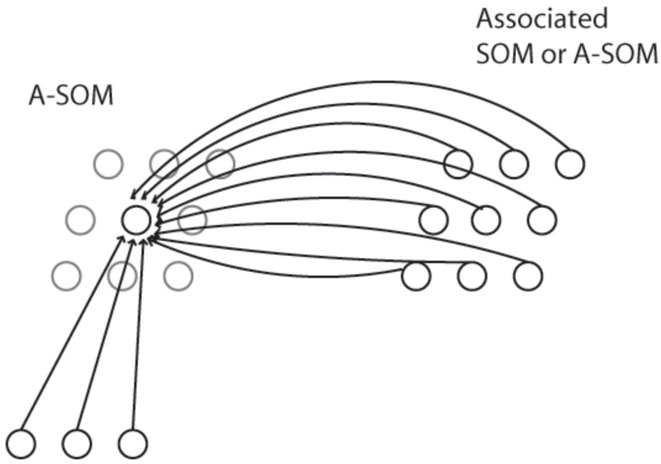


Fig. 1. The connectivity of the A-SOM network. During training each neuron in an A-SOM receives two kinds of input. One kind of input is the native input, which correspond to the input an ordinary SOM receives. The other kind of input is the activity of each neuron in an associated SOM or A-SOM. In the fully trained A-SOM, activity can be triggered by either native input or by activity in the associated SOM or A-SOM, or both.

where $0 \leq \alpha(t) \leq 1$ is the adaptation strength with $\alpha(t) \to 0$ when $t \to \infty$ and the neighbourhood function $G_{ijc}(t)$ is a Gaussian function decreasing with time.

The weights $w^b_{ijl}$ are adapted by

$$w^b_{ijl}(t+1) = w^b_{ijl}(t) + \beta x^b_l(t)\left[y^{input}_{ij}(t) - y^{extern}_{ij}(t)\right] \quad (6)$$

where $\beta$ is the constant adaptation strength.

## III. SENSORS IN THE EXPERIMENT

The system discussed in this paper employs two sensors (Fig. 2) developed at Lund University Cognitive Science (LUCS). One of these sensors is a texture sensor and the other is a hardness sensor.

The texture sensor consists of a capacitor microphone with a tiny metal edge mounted at the end of a moveable lever, which in turn is mounted on an RC servo. When exploring a material the lever is turned by the RC servo, which moves the microphone with the attached metal edge along a curved path in the horizontal plane. This makes the metal edge slide over the explored material, which creates vibrations in the metal edge with frequencies that depend on the textural properties of the material. The vibrations are transferred to the microphone since there is contact between it and the metal edge. The signals are then sampled and digitalized by a NiDaq 6008 (National Instruments) and conveyed to a computer via a USB-port. The FFT is then applied to the input, thus yielding a spectrogram of 2049 component frequencies.

The hardness sensor consists of a stick mounted on a RC servo. During the exploration of a material the RC servo tries to move to a certain position, which causes a downward movement of the connected stick at a constant pressure. In the control circuit inside the RC servo there is a variable

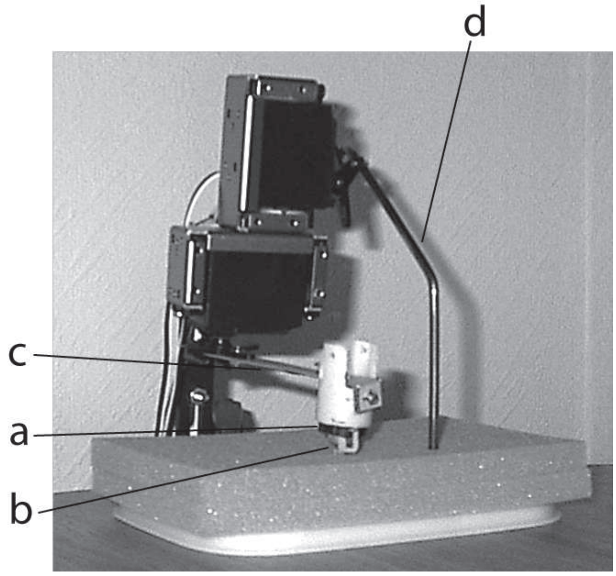


Fig. 2. The texture and hardness sensors while exploring a piece of foam rubber. The texture sensor consists of a capacitor microphone (a) with a metal edge (b) mounted at the end of a moveable lever (c), which in turn is mounted on a RC servo. The hardness sensor consists of a stick (d) mounted on a RC servo. The servo belonging to the hardness sensor contains a variable resistor that provides a measure of the turning of the servo, and thus the displacement of the stick, which is proportional to the compression of the explored material. The actuators are controlled via a SSC-32 controller board (Lynxmotion Inc.). The measure of the resistance of the variable resistor in the RC servo for the hardness sensor and the microphone signal of the texture sensor are digitalized using a NiDaq 6008 (National Instruments) and conveyed to the computer via a USB-port.

resistor that provides the control circuit with information whether the RC servo has reached the wanted position or not. In our design, we measure the value of this variable resistor at the end of the exploration of the material and thus get a measure of the end position of the stick in the exploration. This end position is proportional to the compression of the explored material. The value of the variable resistor is conveyed to a computer and represented in binary form.

The actuators for both the sensors are controlled from the computer via a SSC-32 controller board (Lynxmotion Inc.). The software for the system presented in this paper is developed in C++ and runs within the Ikaros system [1][2][10]. Ikaros provides an infrastructure for computer simulations of the brain and for robot control.

## IV. EXPLORATION OF OBJECTS

The system described in this paper have been trained and tested with two sets of samples. One set consists of 40 samples of texture data and the other set consists of 40 samples of hardness data. These sets have been constructed by letting the sensors explore each of the eight objects described in Table 1 five times.

During the hardness exploration of an object the tip of the hardness sensor stick (Fig. 2d) is pressed against the object with a constant force and the displacement is measured.

The exploration with the texture sensor is done by letting its lever (Fig. 2c) turn 36 degrees during one second. During this movement the vibrations from the metal edge (Fig. 2b) slid over the object are recorded by the microphone (Fig. 2a) mounted at the end of the stick.

The output from the texture sensor from all these explorations has then been written to a file after the application of the FFT. Likewise the output from the hardness sensor has been written to a file represented as binary numbers. The hardness samples can be considered to be binary vectors of length 18 whereas the texture samples can be considered to be vectors of length 2049. The eight objects have various kinds of texture and can be divided into two groups, one with four rather soft objects and one with four rather hard objects. During the exploration, the objects were fixed in the same location under the sensors.

## V. EXPERIMENT

Our system is a bimodal model of haptic hardness and texture perception (Fig. 3). It consists of two monomodal subsystems (hardness and texture), which develop monomodal representations (A-SOMs) that are associated with each other. The subsystem for hardness uses the raw sensor output from the hardness sensor, represented as a binary number with 18 bits and conveys it to an A-SOM with $15 \times 15$ neurons. After training, this A-SOM will represent the hardness property of the explored objects.

In the subsystem for texture, the raw sensor output from the texture sensor is transformed by a FFT module into a spectrogram containing 2049 frequencies, and the spectrogram which is represented by a vector, is in turn conveyed to an A-SOM with $15 \times 15$ neurons. After training, this A-SOM will represent the textural properties of the explored objects.

The two subsystems are coupled to each other in that their A-SOMs also receive their respective activities as associative input.

Both A-SOMs begun their training with the neighbourhood radius equal to $15$. The neighbourhood radius was decreased at each iteration by multiplication with $0.998$ until it reached the minimum neighbourhood size 1. Both A-SOMs started out with $\alpha(0) = 0.1$ and decreased it by multiplication with $0.9999$. $\beta$ where set to $0.35$ for both A-SOMs.

The system was trained with samples from the training set, described in the previous section, by 2000 iterations before evaluation.

## VI. RESULTS AND DISCUSSION

The results of the experiment are depicted in Fig. 4. The 6 images depict the centres of activation when the fully trained system was tested with the test set (described above) constructed with the aid of the objects a-h in Table 1. Images 4A, 4B and 4C correspond to the texture representing A-SOM. Likewise the images 4D, 4E and 4F correspond to the hardness representing A-SOM. Each cell in an image represents a neuron in the A-SOM. In the images 4A, 4B, 4D and 4E there are black circles in some of the cells. This means that the corresponding neurons in the A-SOM are the

TABLE I

THE EIGHT OBJECTS USED IN THE EXPERIMENT. THE OBJECTS A-H WERE USED BOTH FOR TRAINING AND TESTING. THE MATERIALS OF THE OBJECTS ARE PRESENTED AND THEY ARE SUBJECTIVELY CLASSIFIED AS EITHER HARD OR SOFT BY THE AUTHORS. A ROUGH SUBJECTIVE ESTIMATION OF THEIR TEXTURAL PROPERTIES IS ALSO PROVIDED.

| Label | Object | Estimated Hardness | Estimated Texture |
|---|---|---|---|
| a | Foam Rubber | Soft | Somewhat Fine |
| b | Hardcover Book | Hard | Shiny |
| c | Bundle of Paper | Hard | Fine |
| d | Cork Doily | Hard | Rough |
| e | Wood Doily | Hard | Fine |
| f | Bundle of Denim | Soft | Somewhat Fine |
| g | Bundle of Cotton Fabric | Soft | Somewhat Fine |
| h | Terry Cloth Fabric | Soft | Rough |

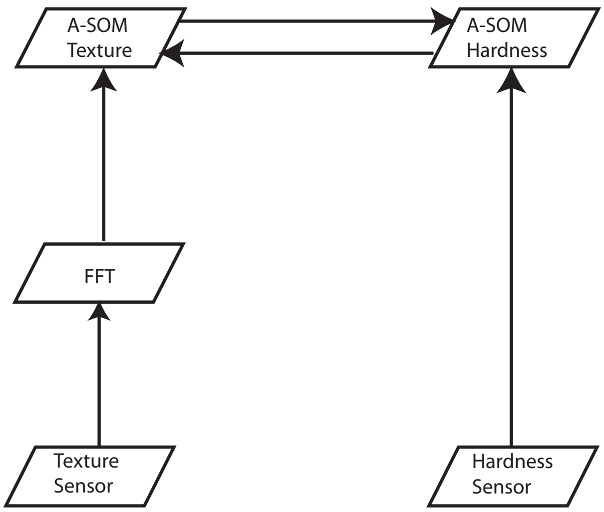


Fig. 3. Schematic depiction over the architecture of the haptic hardness and texture perception system. The system consists of two monomodal subsystems, which develop monomodal representations (A-SOMs) of hardness and texture that learn to associate their activities. The hardness subsystem uses the raw sensor output from the hardness sensor as input to an A-SOM, which finds a representation of the hardness property of the explored objects. The texture subsystem transforms the raw sensory data by the aid of a FFT module and then forwards it to another A-SOM, which finds a representation of the textural properties of the explored objects. The two A-SOMs learn to associate their respective activities.

centre of activation for one or several of the samples in the test set. The centres of activation from the samples in the test set corresponding to each object in Tab 1 when only native input was provided have been encircled in 4A and 4D to show where different objects are mapped in the A-SOMs. Native input should be understood as texture input for the texture representing A-SOM, and hardness input for the hardness representing A-SOM. These results with only native input to the A-SOMs are similar to our earlier results with the hardness and texture sensors together with ordinary SOMs [11]. The encirclings are also present in the other four images. This is so because we want to show how the A-SOMs are activated when there are both native and external input provided to the system (4B and 4E), and when there are only external input provided (4C and 4F). External input should be understood as hardness input in the case of the texture representing A-SOM, and as texture input in the case of the hardness representing A-SOM.

Fig. 4A depicts the texture representing A-SOM in the fully trained system when tested with the test set (only native texture input). As can be seen, most objects are mapped at separate sites in the A-SOM (c, d, e, f, h). There are some exceptions though, namely a, b and g. So the system is able to discriminate between individual objects when provided with native input only, although not perfectly.

The hardness representing A-SOM in the fully trained system when tested with the test set (only native hardness input), depicted in Fig. 4D, also maps different objects at different sites in the A-SOM but not as good as the texture representing A-SOM. The hardness representing A-SOM recognizes b, f and h perfectly and blurs the other more or less. However, the hardness representing A-SOM perfectly discriminates hard from soft objects.

When the texture representing A-SOM receives native texture input as well as external hardness input (as can be seen in Fig. 4B) its activations are very similar to those in Fig. 4A. Likewise when the hardness representing A-SOM receives native hardness input as well as external texture input (as can be seen in Fig. 4E) its activations are very similar to those in Fig. 4D.

Fig. 4C depicts the activations in the texture representing A-SOM when it receives only external hardness input. As can be seen this external hardness input very often triggers an activity similar to the activity following native texture input. Likewise Fig. 4F depicts the activity in the hardness representing A-SOM when it receives only external texture input. Even in this case the external input very often triggers an activity similar to the activity following native input. This means that when just one modality in the system receives input, this can trigger activation in the other modality similar to the activation in that modality when receiving native input. Thus an object explored by both sensors during training of the system can trigger a more or less proper activation in the representations of both modalities even when it can be explored by just one sensor during testing. However, as can be seen in Fig. 4C and Fig. 4F, the activity triggered solely by external input does not map every sample properly. The worst cases are the objects c, d and g in the texture representing A-SOM (Fig. 4C) and the objects a, b and g in the hardness representing A-SOM (Fig. 4D). As can be seen in Fig. 4D, the objects c, d and g are not distinguishable in the hardness representing A-SOM, and the objects a, b and g are not distinguishable in the texture representing A-SOM (Fig. 4A). Thus the external activity patterns for these objects are overlapping and the receiving A-SOM cannot be expected to learn to map these patterns correctly even if the objects where well separated by the A-SOM when it received native input.

## VII. CONCLUSION

We have experimented with a bimodal self-organizing system for object recognition, which is based on textural and hardness input and with associated representations of the two submodalities. The texture sensor employed is based on the transmission of vibrations to a microphone when the sensor slides over the surface of the explored material. The hardness sensor is based on the measurement of displacement of a stick when pressed against the material at a constant pressure. The results are encouraging, both for the developed monomodal representations and the systems ability to associate the activity in these representations. The system is able to discriminate individual objects based on input from each submodality and to discriminate hard from soft objects. In addition input to one submodality can trigger an activation pattern in the other submodality, which resembles the pattern of activity the object would yield if explored with the sensor for this other submodality.

Our experiments with texture complement those done by Edwards et al [4] and Hosoda et al [6] because they only show that the signals from their sensors are in principle useful as texture sensors whereas we actually implement a working self-organizing system. When compared to the work done by Mayol-Cuevas et al [14] our texture experiments differ in that we use a sensor that is not manually rubbed over the material as their pen, but moved by an actuator built into the sensor. A couple of extensions in our experiments when compared to all the previously mentioned experiments and to the work done by Campos and Bajcsy [3] are that we also experimented with both hardness and texture and the association between these two submodalities.

Because of the successful approach to base a texture sensor on a microphone and base hardness perception on the measurements of displacements at a constant applied pressure, we will in the future try to integrate this approach with our haptic systems. We will also continue our experimentations with the A-SOM. We will continue by testing the ability of the A-SOMs when there are very many categories to see if the A-SOM works equally good in that case. We will also try to implement an extended version of the A-SOM, which can be associated with several external A-SOMs or SOMs. In this way we could build multimodal systems, which when receiving input from just one modality would trigger proper activation patterns in the other modalities as well. This extension should be quite straightforward. It could be done by just adding a new weight vector to each neuron for every new associated A-SOM or SOM. The activity of the neurons would be calculated by adding the native activity and the activities coming from all associated A-SOMs or SOMs and divide the sum with the total numbers of activities.

Another very interesting continuation, since we focus much of our research on haptics, would be to test this A-SOM technology in systems that integrate visual and haptic subsystems. In this way we could probably get a visual system to trigger a proper apprehension of a robot hand, in this very bio-inspired way, when it is about to grasp an object.

## REFERENCES


[1] Balkenius, C., & Morén, J. (2003). From isolated components to cognitive systems. *ERCIM News, April 2003*, 16.

[2] Balkenius, C., Morén, J. & Johansson, B. (2007). Building system-level cognitive models with Ikaros. *Lund University Cognitive Studies, 133*.

[3] Campos, M. & Bajcsy, R. (1991). A Robotic Haptic System Architecture, *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, 338-343.

[4] Edwards, J., Melhuish, C., Lawry, J., & Rossiter, J. (2007). Feature Identification for Texture Discrimination from Tactile Sensors. *Proceedings of TAROS 2007*, University of Wales, Aberystwyth, UK, 115-121.

[5] Gardner, E.P., Martin, J.H., & Jessell, T.M. (2000). The bodily senses. In Kandel, E.R., Schwartz, J.H., & Jessell, T.M., (ed.). *Principles of neural science*, 430-450, McGraw-Hill.

[6] Hosoda, K., Tada, Y., & Asada, M. (2006). Anthropomorphic robotic soft fingertip with randomly distributed receptors, *Journal of Robotics and Autonomous Systems*, 54, 2, 104-109.

[7] Johnsson, M., & Balkenius, C. (2006). Experiments with Artificial Haptic Perception in a Robotic Hand, *Journal of Intelligent and Fuzzy Systems*.

[8] Johnsson, M., & Balkenius, C. (2007a). Neural Network Models of Haptic Shape Perception, *Journal of Robotics and Autonomous Systems*, 55, 720-727

[9] Johnsson, M., & Balkenius, C. (2007b). Experiments with Proprioception in a Self-Organizing System for Haptic Perception. *Proceedings of TAROS 2007*, University of Wales, Aberystwyth, UK, 239-245.

[10] Johnsson, M., & Balkenius, C. (2007c). LUCS Haptic Hand III - An Anthropomorphic Robot Hand with Proprioception. *LUCS Minor 13*.

[11] Johnsson, M., & Balkenius, C. (2008). Recognizing Texture and Hardness by Touch. *To appear in the proceedings of IROS 2008*, Nice, France.

[12] Kohonen, T. (1988). *Self-Organization and Associative Memory*, Berlin Heidelberg, Springer-Verlag.

[13] Kohonen, T. (1990). The self-organizing map, *Proceedings of the IEEE*, 78, 9, 1464-1480.

[14] Mayol-Cuevas, W. W., Juarez-Guerrero, J., & Munoz-Gutierrez, S. (1998). A First Approach to Tactile Texture Recognition, *IEEE International Conference on Systems, Man, and Cybernetics*, 5, 4246-4250.

[15] Saper, C.B., Iversen, S., & Frackowiak, R. (2000). Integration of sensory and motor function: The association areas of the cerebral cortex and the cognitive capabilities of the brain. In Kandel, E.R., Schwartz, J.H., & Jessell, T.M., (ed.). *Principles of neural science*, 349-380, McGraw-Hill.

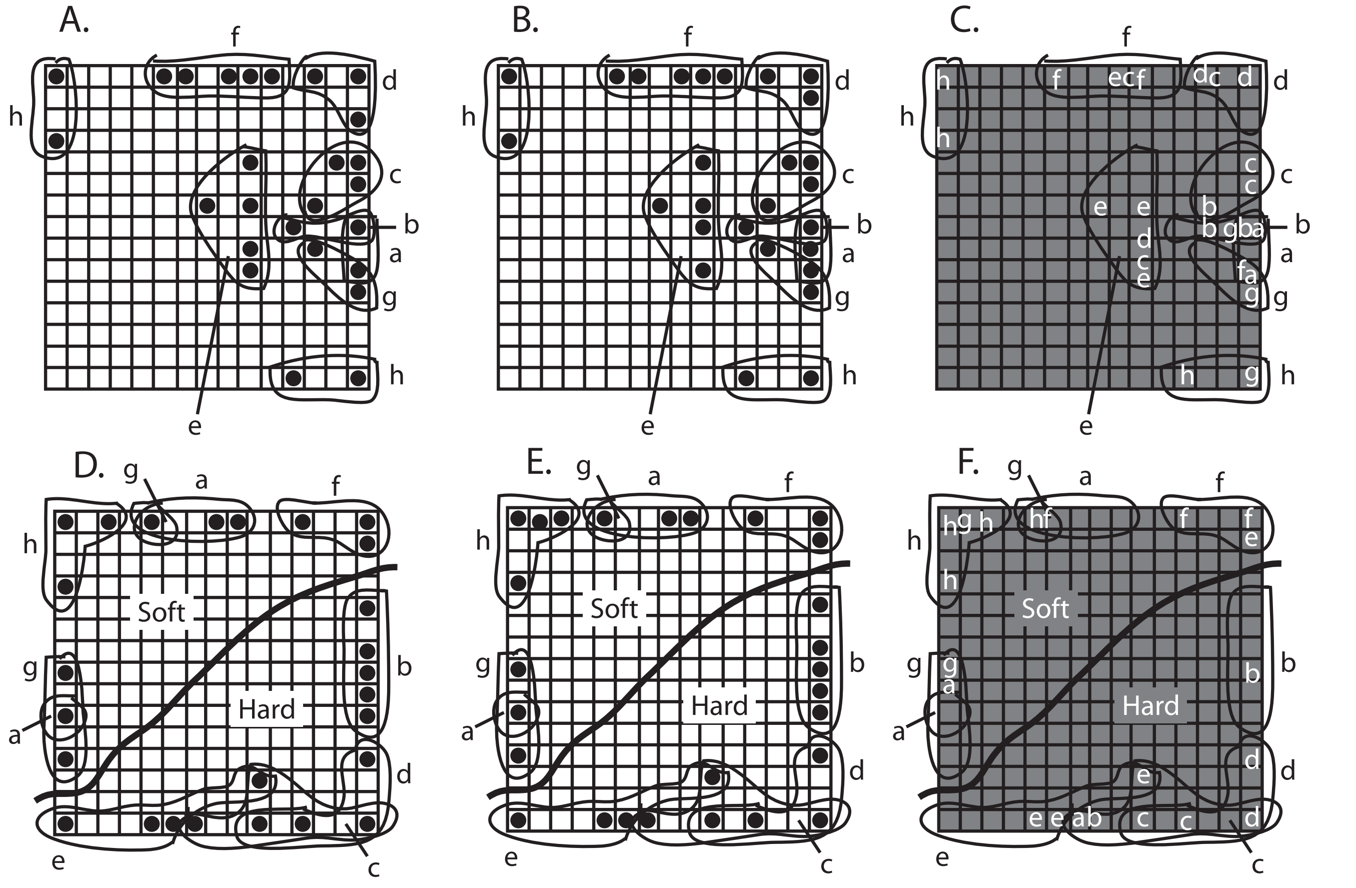


Fig. 4. The mapping of the objects used in the experiments. The characters a-h refer to the objects in Table 1. The images in the uppermost row correspond to the texture representing A-SOM and the images in the lowermost row correspond to the hardness representing A-SOM. Each cell in an image represents a neuron in the A-SOM, which consists of $15 \times 15 = 225$ neurons. A filled circle in a cell is supposed to mean that that particular neuron is the centre of activation for one or several explorations. The occurrence of a certain letter in the rightmost images means that there are one or several centres of activation for that particular object at that particular place. The centres of activation from the samples in the test set corresponding to each object in Tab 1 when only native input was provided have been encircled in the images. A: The texture representing A-SOM when tested with native texture input. Most objects are mapped at separate sites so the system is able to discriminate between individual objects when provided with native input, although not perfectly. B: The texture representing A-SOM when tested with native texture input together with external hardness input. Its activations are very similar to those in A. C: The texture representing A-SOM when it receives only external hardness input. This often triggers an activity similar to the activity following native texture input. D: The hardness representing A-SOM when tested with native hardness input maps different objects at different sites and it perfectly discriminates hard from soft objects. E: The hardness representing A-SOM when tested with native hardness input together with external texture input. Its activations are very similar to those in D. F. the hardness representing A-SOM when it receives only external texture input. This often triggers an activity similar to the activity following native hardness input.