Recognizing Texture and Hardness by Touch

Magnus Johnsson and Christian Balkenius

***Abstract*— We have experimented with different neural network based architectures for bio-inspired self-organizing texture and hardness perception systems. 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. We have implemented and successfully tested both monomodal systems for texture and hardness perception and multimodal systems that merge texture and hardness data into one representation. All systems were trained and tested with multiple samples gained from the exploration of a set of 4 soft and 4 hard objects of different materials. The monomodal texture system was good at mapping individual objects in a sensible way, the hardness systems was good at mapping individual objects and in addition dividing the objects into categories of hard and soft objects. The multimodal system was successful in merging the two modalities into a representation that performed at least as good as the best recognizer of individual objects, i.e. the texture system, and at the same time categorizing the objects into hard and soft.**

I. INTRODUCTION

Two important submodalities in haptic perception are texture and hardness perception. In non-interactive tasks, the estimation of properties like the size and the shape of an external object is often to a large extent based on vision only and haptic perception will only be employed when visual information about the object is not reliable. This might happen for example at bad lighting conditions or when the object is more or less occluded. Haptic submodalities like texture and hardness perception are different in this respect. These submodalities are especially important because they provide information about the outer world that is unavailable for all the other perception channels.

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 in a test enabled the discrimination of five different materials pushed and rubbed by the fingertip. Mayol-Cuevas et al [13] 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). Edwards et al [5] 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 [4] have 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.

Our 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) [11] and a neural network with leaky integrators. A SOM is a self-organizing neural network that finds a low-dimensional discretized and topology preserving representation of the input space. The system 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 neural network based models of texture and hardness perception as well as models that merges

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

these submodalities. All models employ a microphone based texture sensor and/or a hardness sensor that measures the compression of the material at a constant pressure.

II. SENSORS IN THE EXPERIMENTS

All models discussed in this paper employ at least one of two sensors (Fig. 1) 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 then conveyed to a computer via an 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 resistor that provides the control circuit with information whether the RC servo has been reaching 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 all systems presented in this paper is developed in C++ and much of it runs within the Ikaros system [1][2]. Ikaros provides an infrastructure for computer simulations of the brain and for robot control.

III. EXPLORATION OF OBJECTS

Each model described in this paper have been trained and tested with one or both of 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 Tab. 1 five times.

During the hardness exploration of an object the tip of the hardness sensor stick (Fig. 1d) is pressed against the object with a constant force and the displacement where a fixe sensor reading is obtained is measured.

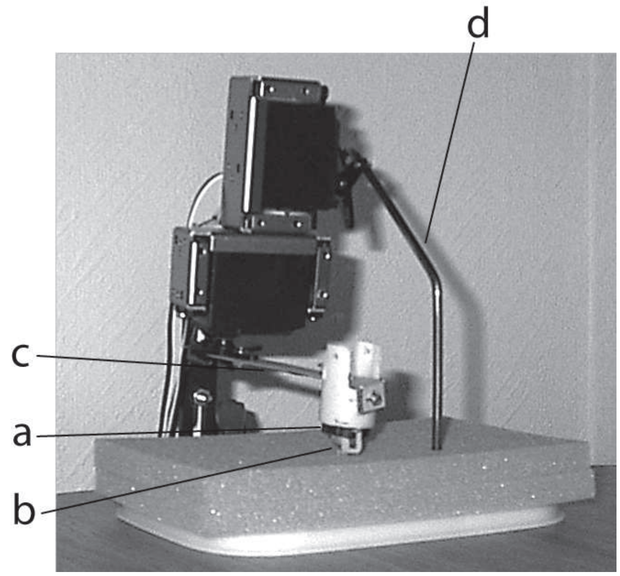

Fig. 1. The texture and hardness sensors while exploring a piece of foam rubber. The texture sensor consists of a capacitor microphone (a) with a metal edge (b) mounted at the end of a moveable lever (c), which in turn is mounted on a RC servo. The hardness sensor consists of a stick (d) mounted on a 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 an USB-port.

The exploration with the texture sensor is done by letting its lever (Fig. 1c) turn 36 degrees during one second. During this movement the vibrations from the metal edge (Fig. 1b) slid over the object are recorded by the microphone (Fig. 1a) 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.

IV. A TEXTURE PERCEPTION MODEL

The texture perception model (Fig. 2 A) is a monomodal model. This means that the raw sensor output from the texture sensor is transformed by the FFT into a spectrogram containing 2049 frequencies, and the spectrogram represented by a vector is in turn conveyed to a SOM, which uses softmax activation [3] with the softmax exponent equal to 10. After training the SOM will represent the textural properties of the explored objects.

TABLE I

THE EIGHT OBJECTS USED IN THE EXPERIMENTS WITH THE TEXTURE AND THE HARDNESS MODELS. 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. 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 Jeans Fabric	Soft	Somewhat Fine
g	Bundle of Cotton Fabric	Soft	Somewhat Fine
h	Terry Cloth Fabric	Soft	Rough

We have experimented with different parameter settings of the texture SOM, both with the aim to get a well-working monomodal model and to get a model that would serve well as a part of a multimodal model, and we reached the conclusion that a well-working set of parameters is to use a SOM with 15×15 neurons with a plane topology. A torus topology was also tested but turned out to be less effective than a plane topology. The sort of topology used influences the behaviour of the SOM at the borders. With plane topology the activations from the objects in the training set tend to be close to the borders, which turned out to be good when the texture perception model was used as a part of the combined monomodal/multimodal model described below. We also experimented with different decay rates of the Gaussian neighbourhood function. We came to the conclusion that a neighbourhood radius of 15 at the start of the training phase, which decreased gradually until it was approximately 1 after 1000 iterations, and stayed at 1 during the rest of the training phase, was satisfactory. This model and all the others were trained during 2000 iterations before evaluation. We reasoned that it would be good if the neighbourhood had shrunk to a small value after about 1000 iterations in order to let the multimodal SOM of the combined model, described below, get enough iterations to self-organize. In other words, the idea was that the texture SOM should be rather well organized after 1000 iterations.

V. A HARDNESS PERCEPTION MODEL

The hardness perception model is also monomodal. In this model (Fig. 2 B), the raw sensor output from the hardness sensor, represented as a binary number with 18 bits, is conveyed to a SOM, which like the texture model uses softmax activation with the softmax exponent equal to 10. After training, the SOM will represent the hardness property of the explored objects.

As in the case of the texture model we have experimented with different parameter settings of the hardness SOM, and for the same reasons. In this case we also tested a lot of different sizes of the monomodal SOM. This was because preliminary experiments indicated that it could be a good idea to use a very small sized SOM for hardness in the combined model described below. This was because a small sized hardness SOM seemed to self-organize solely according to

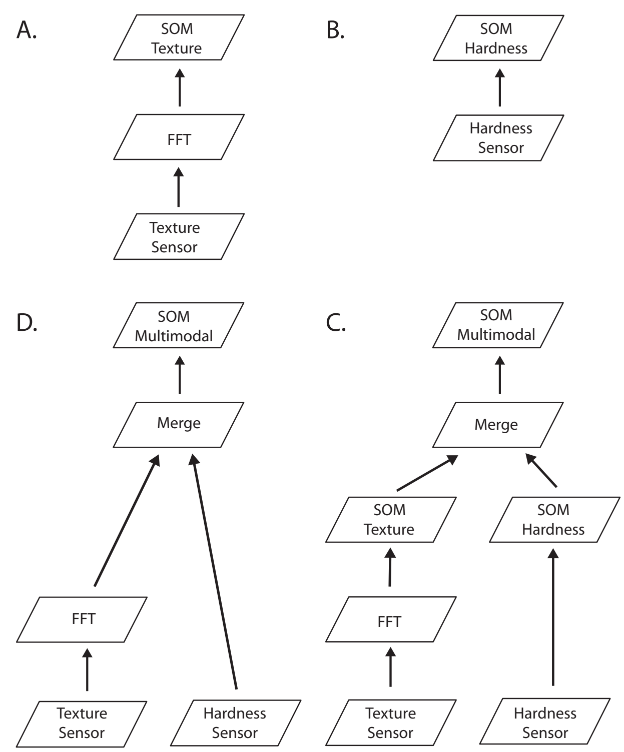

Fig. 2. Schematic depiction of the good working model architectures. A: A monomodal model of texture perception. The raw sensor output is transformed by the FFT into a spectrogram containing 2049 frequencies. The spectrogram represented by a vector is conveyed to a SOM. B: A monomodal model of hardness perception. The raw sensor output represented as a binary vector with 18 elements is conveyed to a SOM. C: A model with both monomodal and multimodal representations. This model could be seen as a merging and an extension of the previous two models, or likewise the previous two models could be seen as the monomodal level of this model. The output from the texture SOM and the output from the hardness SOM is merged, i.e. a new vector is created by transforming the activations of the texture SOM and the hardness SOM into vectors and putting them after each other. The merged vector is used as input to a multimodal SOM. This means that in this model there are self-organizing representations of texture and hardness as well as a combined representation of both. D: A multimodal model. This model directly combines the output from the FFT and the binary output from the hardness sensor into a new vector in the same way as described in the previous model, but without the step with monomodal representations. The combined vector is used as input to a multimodal SOM.

the hardness property and not distinguish individual objects, and since the texture SOM was better at distinguishing individual objects we did not want the hardness part to blur this although we wanted it to make the multimodal representation become organized according to hardness as well. We tried SOMs with planar as well as torus topology and with 15×15, 10×10, 5×5, 2×2 or 1×2 neurons. All variants started with a neighbourhood size that covered the whole SOM and the rates of decay of the neighbourhood were adjusted so that the neighbourhood would shrink to a radius of approximately 1 after about 1000 iterations. As we had expected the 15×15 neurons SOM (with plane topology) was best in this monomodal model but we also found that, as suspected, all tested sizes but one indeed organized to divide the objects into the categories hard and soft ones. The exception was the SOM with only 1×2 neurons, which did not preserve the division of hard and soft objects in a good way.

VI. A COMBINED MONOMODAL AND MULTIMODAL MODEL

In this model (Fig. 2 C) we experimented with different ways of combining the output from the monomodal SOMs to an input for the multimodal SOM. First we tried a method for cross coding that we have used in other contexts. In this method a two-vector input self-organizing neural network called T-MPSOM [8] that self-organize into something similar to tensor product operation, which is an operation on two vectors resulting in a matrix, was used to combine the outputs from the monomodal SOMs. In previous research the T-MPSOM was very successful in coding proprioceptive information and it also worked in the current model. However, we also experimented with a simpler method of combining the monomodal outputs, which was also superior for this aim. This method was simply to combine the activity of the monomodal SOMs, re-arranged into vectors, by creating a new vector by putting the hardness output vector after the texture output vector.

The monomodal texture SOM used 15×15 neurons with the same parameter setting as in the texture model. In the case of the monomodal hardness SOM we tried two different variations, namely a 2×2 neurons SOM and a 15×15 neurons SOM with the settings specified in the hardness model above. Both worked fine but the variation with the 2×2 neurons SOM yielded the best representation in the multimodal SOM. The multimodal SOM had similar settings as the monomodal texture SOM, but the decay rate of the neighbourhood was set to decrease the neighbourhood radius to one in 2000 iterations.

VII. A MULTIMODAL MODEL

In the multimodal model (Fig. 2 D) we combined the output from the texture sensor, after transformation into a spectrogram by a FFT, with the raw hardness sensor output expressed as a binary number by the same method as in the combined model described above, i.e. by putting the output vector from the hardness sensor after the output vector from the FFT. This means that this model has no monomodal representations. The combined vector was used as input to a multimodal SOM with the same settings as in the combined model above. Also in this model we tried to use T-MPSOM but with a worse result than with this simpler method.

VIII. RESULTS AND DISCUSSION

The mapping of the objects (a-h in Tab. 1) used in the experiments with the different models is depicted in Fig. 3. Each image in the figure corresponds to a SOM in a fully trained model and each cell in an image corresponds to a neuron in the corresponding SOM. A filled circle in a cell is supposed to mean that that particular neuron is the centre of activation in one or several explorations. In fig. 3A the mapping of individual texture explorations with the texture model have been encircled. As can be seen, most objects are mapped at separate sites in the SOM (c, d, e, f, h). There are some exceptions though, namely a, b and g. So the texture model is able to discriminate between individual objects, although not perfectly.

The SOM in the hardness model, depicted in Fig. 3 B, also maps different objects at different sites in the SOM but not as good as the texture model. The hardness model recognizes b, f and h perfectly and blurs the other more or less. However, the model perfectly discriminates hard from soft objects.

The combined monomodal and multimodal model (Fig. 3 C), which as mentioned above can be seen as a merging and extension of the texture model and the hardness model (with 2×2 neurons in the SOM), discriminate hard from soft objects well. In two explorations the hard/soft category is undetermined. This is so because one exploration of an object a and one exploration of an object g have the same centre of activation. It also discriminates perfectly between the objects b, d, f and h.

The multimodal model (Fig. 3 D) discriminates perfectly between the objects c, d, e, f and h, i.e. the same objects as in the texture model. Moreover, it also discriminates hard from soft objects, although in seven explorations the hard/soft category is undetermined because three explorations of the object a and four explorations of the object b have the same centre of activation.

An interesting observation can be made from Fig. 3, namely that objects mapped close to each other in the texture model also tend to be mapped close to each other in the combined monomodal and multimodal model and in the multimodal model, but not so in the hardness model. This could be interpreted as that the multimodal SOMs seek to preserve the texture map but re-organize it so that hard and soft objects are discriminated.

Our experiments with texture complement those done by Edwards et al [5] 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 system. When compared to the work done by Mayol-Cuevas et al [13] 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

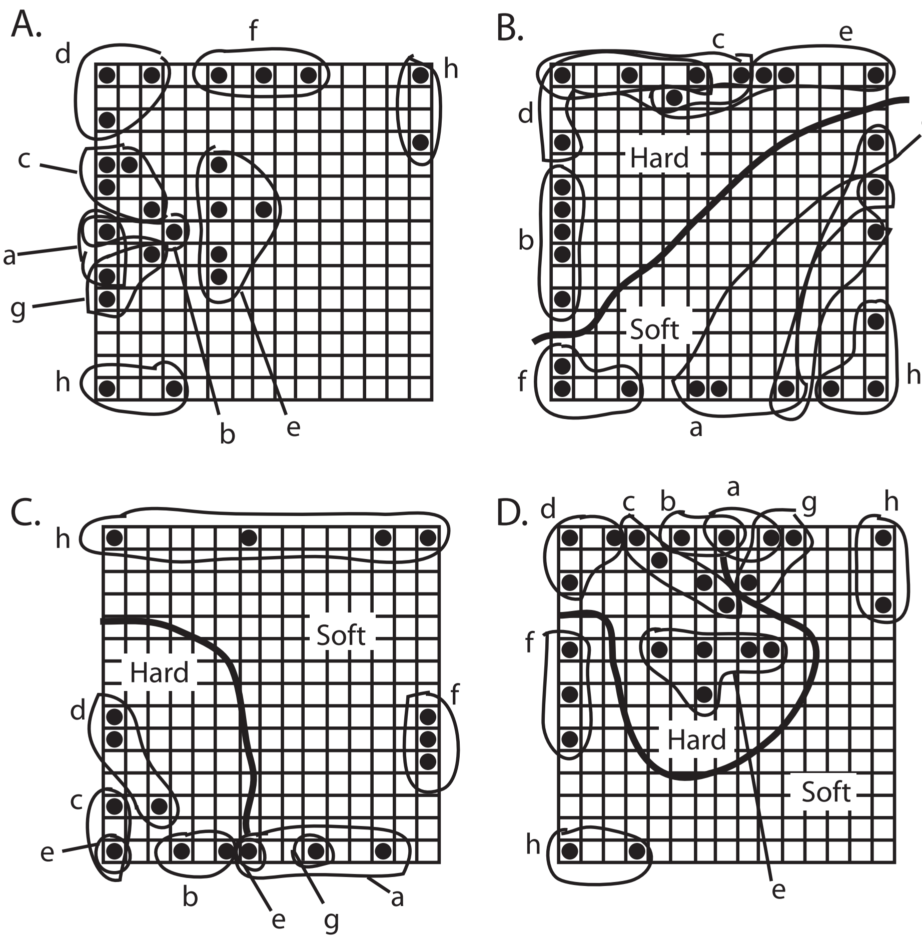

Fig. 3. The mapping of the objects used in the experiments. The characters a-h refer to the objects in Tab. 1. Each image in the figure corresponds to a SOM in a fully trained model and each square represents a neuron in the SOM, which consists of $15 \times 15 = 225$ neurons. A filled circle in a cell is supposed to mean that that particular neuron is the centre of activation for one or several explorations. The occurrence of a certain letter at more than one place means that the corresponding object has different centres of activation during different explorations of the same object, i.e. all letters of a certain kind represents all occurring centres of activation in the SOM when the system was tested with the corresponding object. A: The monomodal SOM in the texture model. The centres of activation of all instances of each object have been encircled. The objects c, d, e, f and h are mapped at non-overlapping sites in the SOM, whereas the objects a, b and g are not. This can be interpreted as that the texture model is able to discriminate between individual objects, although not perfectly. B: The monomodal SOM in the hardness model. In this model the objects b, f and h are perfectly recognized, whereas the others are not. Moreover, the model perfectly discriminates hard from soft objects. C: The multimodal SOM in the combined monomodal and multimodal model. In this model the objects b, d, f and h are perfectly recognized, whereas the others are not. It also discriminates hard from soft objects. In two explorations the hard/soft category is undetermined, because one exploration of an object a and one exploration of an object g have the same centre of activation. D: The multimodal SOM in the multimodal model. In this model the objects c, d, e, f and h are perfectly discriminated, i.e. the same objects as in the texture model. Moreover, it also discriminates hard from soft objects, although in seven explorations the hard/soft category is undetermined because three explorations of the object a and four explorations of the object b have the same centre of activation.

the sensor. Another difference is that our system is self-organizing. An extension in our experiments when compared to all the previously mentioned experiments and to the work done by Campos and Bajcsy [4] is that we also experimented with both hardness and texture and the merging of these two submodalities. Like the work of Mazid and Russel [12], our system detects the profile of a surface, but is based on the much simpler principle of vibration.

Our work is probably most similar to the system by Takamuku et al [14], although their sensors are rather different. Instead of using microphone based sensors they used a sensor material based on strain gauges within silicone rubber. For exploration they used squeezing as well as tapping. Their work is also similar in that self-organizing maps were used for each submodality. However, this robotic system did not investigate the use of multimodal convergence which is the main contribution of our work.

In conclusion the multimodal model seems to be the best one since it recognizes as many individual objects as the monomodal texture model and in addition recognizes most of the objects as either soft or hard. However the combined monomodal and multimodal model recognizes the hardness of more objects correctly, but it comes at the price of being slightly worse at recognizing individual objects.

IX. CONCLUSION

We have experimented with several self-organizing systems for object recognition based on textural and/or hardness input. 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 monomodal systems and the multimodal systems. The multimodal systems seem to benefit from both submodalities and yield representations that are better than those in the monomodal systems. This is particularly true because the multimodal representations preserve the discrimination ability of the monomodal texture model and also seem to preserve the way that model groups the objects. The influence of the hardness input makes the multimodal representation organize according to hardness as well.

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. In other words we will carry out experiments in which we equip future robot hands with microphone based texture sensors and measure hardness by letting a finger press on the explored material at a constant pressure while measuring the displacement. This could result in systems that explore objects and more or less immediately gain information about the objects shape, size, hardness and textural properties. This will yield a system that is able to discriminate between equally shaped and sized objects made of different materials.

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] Bishop., C. M. (1995). Neural Networks for Pattern Recognition. *Oxford University Press*, Oxford, New York.

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

[5] 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.

[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. (2006a). 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] Kohonen, T. (1988). *Self-Organization and Associative Memory*, Berlin Heidelberg, Springer-Verlag.

[12] Mazid, A.M. & Russell, R.A. (2006). A Robotic Opto-tactile Sensor for Assessing Object Surface Texture. IEEE Conference on Robotics, Automation and Mechatronics, 2006, 1 - 5.

[13] 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.

[14] Takamuku, S., Gómez, G., Hosoda, K. and Pfeifer, E. (2007). Haptic discrimination of material properties by a robotic hand, in: Proc. of the 6th IEEE Int. Conf. on Development and Learning.