A Haptic Texture Database for Tool-mediated Texture Recognition and Classification

Matti Strese*, Jun-Yong Lee†, Clemens Schuwerk*, Qingfu Han†, Hyoung-Gook Kim† and Eckehard Steinbach*
*Chair of Media Technology
Technische Universität München, Germany
Email: matti-strese@mytum.de, {clemens.schuwerk, eckehard.steinbach}@tum.de
†Institute for Multimedia Signal Processing
Kwangwoon University, Republic of Korea
Email: {jasonlee88, qingfu33, hkim}@kw.ac.kr

***Abstract*—While stroking a rigid tool over an object surface, vibrations induced on the tool, which represent the interaction between the tool and the surface texture, can be measured by means of an accelerometer. Such acceleration signals can be used to recognize or to classify object surface textures. The temporal and spectral properties of the acquired signals, however, heavily depend on different parameters like the applied force on the surface or the lateral velocity during the exploration. Robust features that are invariant against such scan-time parameters are currently lacking, but would enable texture classification and recognition using uncontrolled human exploratory movements. In this paper, we introduce a haptic texture database which allows for a systematic analysis of feature candidates. The publicly available database includes recorded accelerations measured during controlled and well-defined texture scans, as well as uncontrolled human free hand texture explorations for 43 different textures. As a preliminary feature analysis, we test and compare six well-established features from audio and speech recognition together with a Gaussian Mixture Model-based classifier on our recorded free hand signals. Among the tested features, best results are achieved using Mel-Frequency Cepstral Coefficients (MFCCs), leading to a texture recognition accuracy of 80.2%.**

I. Introduction

Human haptic perception is composed of the kinesthetic and tactile submodalities, which are both crucial for the perception of objects and surfaces. During the exploration of objects, kinesthetic percepts are important to determine the geometric properties, e.g. the size or the weight of an object. The tactile modality is particularly important to characterize the object's surface [1]. Human-machine or machine-machine interfaces incorporating the haptic modality, thus, have a large potential to increase the technical system's realism and performance [2]. Kinesthetic haptic devices are used to display low-frequency force feedback to the human, commonly mediated through a tool or stylus held by the user. In a teleoperation system, for example, this allows for the physical interaction with the remote environment. During tapping on or stroking of an object's surface with such a tool-mediated interface, high-frequency contact vibrations are produced. Traditional kinesthetic haptic devices, however, are not suitable for displaying those vibrations due to various physical limitations [3], leading to a loss of the rich feel of the object's surface properties. To overcome this limitation, acceleration sensors can be employed to measure the high-frequency contact vibrations (denoted as texture signals in the following), which are then displayed using high-bandwidth bidirectional actuators [4]. Such a vibrotactile display system using measured contact acceleration signals in addition to traditional kinesthetic force feedback, as illustrated in Fig. 1, significantly increases the realism of teleoperation systems [3].

Contact acceleration signals are used in robotic systems to recognize or classify object surface textures, as reviewed in Section II. The robot strokes a tool (or a robotic finger) over an object's surface while a sensor measures the contact vibrations induced on the tool. The main challenge in the texture recognition is the variance in the acceleration signal for a single texture in the time and frequency domain depending on scan-parameters like applied contact force and the scan-velocity. Contrary to audio, speech and image recognition, where well-established features that are invariant against external influences like, e.g. noise or scale, exist, such highly distinctive fingerprints for haptic texture recognition are still lacking. Well-defined exploratory movements with fixed scan-parameters together with extensive training of machine learning algorithms with a variety of different force/velocity combinations are used in robotic texture recognition to overcome this issue (e.g. in [6], [7], [8], [9]).

In a teleoperation system as shown in Fig. 1, however, the human user explores the remote surface texture by controlling the movement of the slave robot. Inherently, scan-parameters like velocity and force are varying and well-controlled surface exploration as used in robotic surface recognition is not possible. Hence, our long-term goal is to identify features providing distinctive fingerprints for texture signals which are invariant

Fig. 1. Teleoperation setup with a vibrotactile display based on a accelerometer and a tactile actuator (adopted from [5]).

978-1-4799-5964-8/14/$31.00 ©2014 IEEE

against different scan-parameters. The benefits of such invariant features are twofold. First, as they rely only on the contact acceleration signals, the usage of additional expensive sensors, e.g. for measuring the contact force, is not necessary. Second, extensive training of machine learning algorithms with many defined scan-parameters is not necessary. Such invariant features, thus, are also potentially benefiting research in robotic surface texture recognition.

As a first step towards the identification of scan-invariant features, it is necessary to establish a database of recorded texture signals representing a set of different texture surfaces (e.g paper, plastic, wood, stone, etc.). Note that the haptics group at the University of Pennsylvania has provided pioneering work with the *Penn Haptic Texture Toolkit (HaTT)*, an open-source collection of contact acceleration signals measured from 100 haptic textures, as well as software algorithms to render these textures using impedance-type haptic devices [10]. While initially working with this database, we recognized the need of having acceleration signals measured during both controlled (known scan-parameters) and uncontrolled (free hand) surface exploration. The HaTT, however, provides only free hand exploration signals with varying force and velocity. Although our objective is to use exactly such free hand recordings, the controlled conditions are more suitable to analyze feature candidates for texture recognition/classification. To start with, we recorded texture signals with controlled and uncontrolled conditions for 43 different textures. The hardware setup, the recording process and the applied signal processing are explained in Section III. We make our database publicly available for download [11]. Note that the purpose of the provided recordings is especially the identification of scan-invariant features. This is in contrast to the HaTT, where the focus is on the rendering and display of textures using texture models similar to the models proposed in [12], [13].

In our previous work in [5] we observed that segments of texture signals show similarities to speech signals and proposed a low-bitrate vibrotactile texture codec for teleoperation systems based on speech coding techniques. Consequently, we adopt well-established features and machine learning algorithms from speech and audio recognition to test them on our new dataset. To the best of our knowledge, there is currently no systematic analysis of the performance of such audio-inspired features for haptic texture recognition. The selected features, the machine learning approach and the recognition results are presented in Section IV.

II. Related work in texture recognition

Contact acceleration signals are often used for robotic texture recognition. An artificial finger mounted on a robotic hand is stroked over a textured surface in [7]. Peaks in the spectral envelope of the measured vibration signal are used as features for material classification using machine learning algorithms. This feature definition depends on the exploration movement as the spectral shape of the signals changes with the scan-parameters [8]. Therefore, extensive training of the classifier with different scan-velocities is necessary while the applied force is always controlled to be constant. The current scan-velocity is then used as an additional feature for the classification. In [9], a three-axis accelerometer is attached to a rigid tool held by a PR2 robot. A single small time frame is extracted from the midpoint of the measured acceleration signal during the exploration movement, its spectrum is calculated and then divided into 30 bins. Those bins, together with the scan-velocity and applied normal force are used as features. Several texture scans covering a large range of different controlled force/velocity combinations are used to train the haptic surface recognition system. Similarly, in [6] the accelerometer is mounted on an artificial fingernail and supplies acceleration signals for texture recognition. Five different but well-defined exploratory movements are used for training and testing of different machine learning algorithms. In [6], the texture signal is first divided into five segments and then the spectrum of each segment is binned into 25 bins. This feature vector represents a spectrotemporal histogram of the texture signal. Other related work presented in [8] uses a biologically inspired robotic finger (the BioTac) together with features describing perceptual dimensions of textures, like e.g. roughness or fineness for texture recognition. [8] identifies a set of useful force/velocity combinations for texture exploration and proposes an algorithm that adaptively selects the next exploration movement. The knowledge gained from several explorations is then fused to recognize a texture.

To summarize, previous approaches in the area of robotic texture recognition mainly use well-defined exploration strategies together with measured scan-parameters as features for texture recognition. In contrast, in our work, we investigate texture recognition and classification using acceleration signals recorded during free hand exploration of haptic textures.

III. Texture Database

A. Hardware setup

The hardware setup we use for recording texture signals is similar to the system proposed in [3] and the one we used in [5]. It is based on a Geomagic Touch haptic device (3DSystems, USA) with a three-axis accelerometer (LIS344ALH from ST Microelectronics) mounted on the stylus of the device as shown in Fig. 2. This MEMS-based analog accelerometer has a range of $\pm$ 6 g.

The stylus of the Geomagic Touch haptic device is equipped by default with a spiky conical aluminum tip. We replace this tip with a customized hemispherical and slightly flattened tip shown in Fig. 2 (diameter of 5 mm), similar to the hemispherical tip used in [14]. Only the tip is in contact with the texture while stroking the surface and may wear off during recording of signals for very rough textures. For this reason, we chose to build the tip from stainless steel instead of aluminum. We decided to use a slightly flattened tip instead of a conical tip because we recognized that the original spiky tip leads to unnatural discontinuities in the exploration of the surface of very coarse textures. This is because the tip gets stuck in the macroscopic structure of the surface, leading to unwanted stick-slip effects. Please note that the shape and size of the tip influences the recorded texture signals, as also discussed in [14]. We plan to extend the database with recordings using different tips to investigate their influence on feature candidates for a texture recognition system in future work.

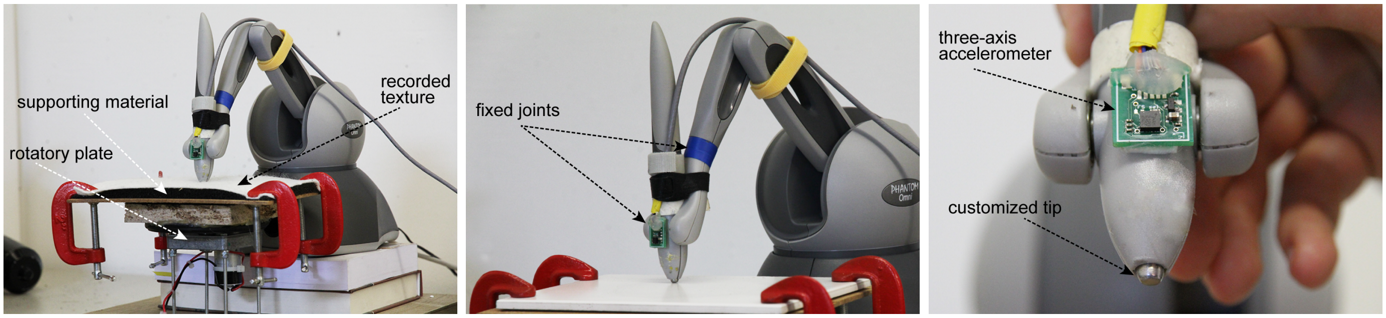

Fig. 2. Hardware setup used for the texture recording. Left: complete setup with the rotary plate used to record the controlled signals. Center: The two marked joints are fixed to ensure upright posture of the stylus. Right: close-up of the customized stainless steel tool tip.

B. Signal processing

A quad-core 2.13 GHz PC running Ubuntu with RTAI real-time extension, together with a NI PCI 6221 data acquisition card is used to sample the accelerometer output at 10 kHz with a 16-bit ADC. When displaying the recorded acceleration signal with a vibrotactile actuator as explained in Section I, the direction of the displayed vibration is not important [15]. For this reason, we use the DFT321 algorithm proposed by Landin *et al.* in [15], which is also used, e.g., in [5], [12], [13] to reduce the dimensionality of the recorded signal. The DFT321 algorithm combines the three-axis acceleration signal into a one-dimensional signal while preserving important temporal and spectral properties.

Afterwards, the texture signal is filtered with a cutoff frequency of 1400 Hz, similar to [9], [10]. This is in contrast to [6] where the contact signals are initially sampled only at 400 Hz. As stated in [3], the increased signal bandwidth is important as humans can perceive texture-induced vibrations of up to 1 kHz. In [9] and [10] a high-pass filter with a cutoff frequency of 10 Hz, respectively 20 Hz, is used to remove the effects of gravity and human motion on the signals. We decided to keep the low frequencies in the signals and to select the cutoff frequency above perceptual limitations as they might still contribute to recognition accuracy. Post-processing of the provided signals is, of course, still possible. Furthermore, the signal is saturated with thresholds of $\pm$ 6 g corresponding to the sensor's output range to secure the validity of the output levels. The filtered signal is then stored in a text file and can be used for further processing, e.g. for feature extraction and texture recognition as explained in Section IV.

C. Textures

Fig. 3 illustrates the set of recorded textures selected from everyday materials. They cover a wide range of material classes, e.g. stone, wood, paper, textile, metal, etc., except for liquids and very sticky surfaces. We also intend to represent a variety of different perceptual dimensions like for example roughness or hardness. Soft surfaces, however, tend to represent the hardness characteristics of the material they are placed on during exploration (e.g. a wooden table). We select a perceptually semi-soft material (medium coarse foam mesh in Fig. 3) as a common supporting pad for all soft surfaces. Our preliminary tests show that the hardness of a piece of leather and a piece of paper can be distinguished using such a supporting surface, while their hardness is perceptually similar if a very hard wooden support is used.

D. Recording process

Two general sets of data have been recorded for this database, which are explained in the following.

1) Controlled recordings: Different scan parameters during the recording of haptic textures cause changes in the temporal and spectral properties of the signals [8], [13]. The exploration velocity, the applied force, the aforementioned tip mounted on the stylus, its shape, or the inclination with which the stylus is stroked over the surface may heavily influence signal properties and, thus, also extracted features used for recognition or classification. As explained in Section II, additional sensors to measure the applied force and scan velocity and well-controlled exploration procedures are often used in robotic texture recognition to compensate for this. The first set of recorded texture signals is based on the setup explained above and adopts controlled conditions for velocity, force and inclination. The motivation for this is to establish a dataset, which can be used to systematically investigate feature candidates and their dependency on velocity and force. We record two different signals for each texture, where either the force is constant and the velocity is increasing over time or vice versa. We adopt a customized rotary plate driven by a propulsive motor to control the velocity of the movement as shown in Fig. 2. Hence, instead of stroking the stylus with controlled velocity over the surface, which is difficult to achieve with our setup without explicit sensors to measure the lateral velocity, the surface itself rotates to simulate the movement of the stylus. The voltage supply for the motor is increased step-wise in 1 V steps to increase the velocity. At the same time, the haptic device is controlled to a fixed position using a PID controller. For the second recording, the voltage supply is kept constant, while the applied force on the texture is linearly increased by changing the target z-position of the PID controller. The recorded texture signal for the material "Plastic Mesh" is shown exemplarily in Fig. 4. We show only the recording with increasing velocity, as the complementary part with increasing force is visually quite similar.

2) Free hand recordings: The second set of texture signals is recorded using less controlled conditions. It consists of ten different free hand recordings (at least 20 s each) conducted by a single human subject. We adopt five lateral movements, where the user moves back and forth between two reversal

Fig. 3. The different textures used to build the database. The textures can be roughly grouped into nine different material categories (1: Stone Types, 2: Glasses, 3: Metalls, 4: Timbers, 5: Rubbers, 6: Plastics, 7: Papers, 8: Foams, 9: Textiles, 10: Floor pads, 11: Leather). Larger pictures of the surfaces are available online [11].

points, and five circular movements on the surface. This is motivated by typical human exploratory movement behavior for object recognition [16]. The scan-parameters and exploratory trajectories are varied between the ten recordings to cover a wide range of typical human texture explorations. Inherently, as a human performs the recordings, also the inclination angle, the applied force as well as the scan-velocity varies within each recording. The motivation to record such free hand recordings is to resemble a real world scenario similar to the teleoperation system shown in Fig. 1, where the human user has full control over the exploration movement. The free hand recordings, thus, provide test and training data for the evaluation of features identified with the help of the controlled recordings.

All recorded texture signals are available for download on our webpage [11], where we also provide more images and videos to illustrate the recording process.

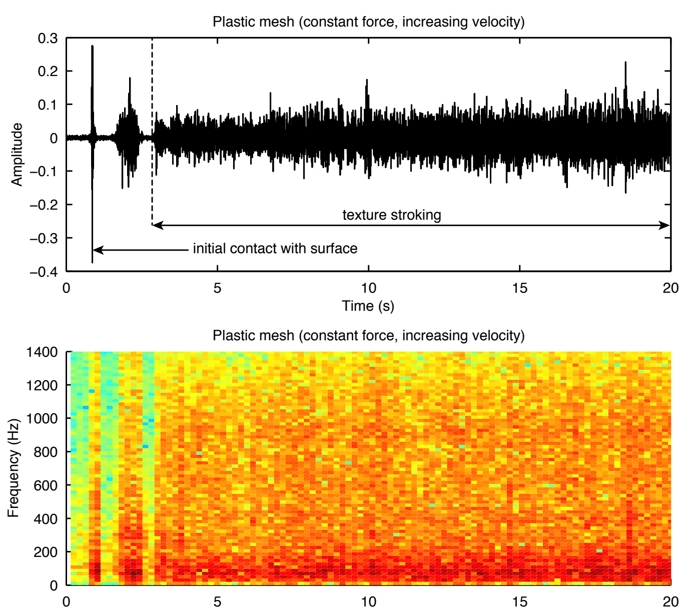

Fig. 4. Recorded texture signal and its spectogram as an example. The amplitude of the measured acceleration increases with increasing velocity. All texture signals include initial contact with the surface, the lateral acceleration to overcome static friction and the actual stroking of the surface.

IV. Texture Classification

As already mentioned in Section I, segments of the recorded texture signals show similarities to speech signals. Therefore, we adopt well-established features from audio content analysis to evaluate their performance in a haptic texture recognition system.

The system consists of three steps: segmentation, feature extraction and texture recognition. First, a segmentation stage isolates relevant texture segments from the accelerometer data traces into three phases - impact on surface, first detected begin of movement and general movement phase. Second, a feature extraction stage extracts a distinctive and low complexity fingerprint for the subsequent texture recognition step. A trade-off between reduced data dimensionality and information content must be found, e.g. by using the dimension reduction approach in [18], which maps the feature vector onto another feature vector with lower dimensionality. Third, a classifier uses the reduced feature vector to recognize the texture. Examples of such classifiers include Gaussian Mixture Models (GMM) [19], Hidden-Markov-Models (HMM) [20], Neural Networks (NN) [21] and Support Vector Machines (SVM) [22]. Compared to NN and SVM, GMM and HMM are based on statistical models. The GMM is independent of the time sequence, while the HMM explicitly models the order of small time frames extracted from the original signal through the transition states. Contrary to audio signals where the temporal order is very meaningful, texture signals do not show such temporal dependencies. We use GMM for texture recognition as explained in Section IV-B since it is the more reasonable statistical approach regarding the presence of only numerical data points by using their centroids and spatial variance orientation in the feature space to build a classification model.

A. Feature Extraction

We test the following seven features commonly used for audio/speech analysis in our texture recognition system.

1) Spectrum Flatness (SF) [18]: Spectral flatness is a measure of the deviation between the signal's spectral form and a flat spectrum. Flat spectra correspond to noise or impulse-like signals, thus, high flatness values indicate noisiness. Low flatness values generally indicate the presence of harmonic components.

2) Spectrum Centroid (SC) [18]: The spectral centroid is defined as the center of gravity of the magnitude spectrum of the short-time Fourier transform (STFT) and, hence, is a measure of the spectral shape. Higher centroids correspond to "brighter" textures including higher frequencies.

3) Linear Predictive Cepstral Coefficient (LPCC) [23]: The Linear Predictive Coefficient (LPC) is one of the most powerful speech analysis techniques and it has gained popularity as a formant estimation technique. The Linear Predictive Cepstral Coefficients (LPCC) are derived from LPCs. LPCCs are more robust against sudden signal changes or noise because they are derived from the impulse response of the speech model.

4) Mel-Frequency Cepstral Coefficient (MFCC) [18]: The mel-frequency cepstral coefficients (MFCC), which are based on the STFT, are widely used as audio features in automatic speech recognition. The power spectrum bins are grouped and smoothed according to the perceptually motivated mel-frequency scaling. The spectrum is segmented into a number of critical bands by means of a filter bank that typically consists of overlapping triangular filters. Finally, a discrete cosine transform applied to the logarithm of the filter bank outputs results in vectors of decorrelated MFCC features.

5) Rasta Perceptual Linear Prediction (RPLP) [24]: The word RASTA stands for *RelAtiveSpecTrAl* Technique. It is an improvement of the Perceptual Linear Prediction (PLP) method and incorporates a special filtering of the different frequency channels of a PLP analyzer. PLP is identical to LPC except that its spectral characteristics have been transformed to match characteristics of the human auditory system.

6) Chromagram (CG) [25]: Chroma-based audio features have turned out to be a powerful tool for various analysis tasks in music information retrieval, including tasks such as chord labeling, music summarization, structure analysis, music synchronization and audio alignment. A 12-dimensional chroma feature encodes the short-time energy distribution of the underlying music signals over the twelve chroma bands, which correspond to the twelve traditional pitch classes of the equal-tempered scale encoded by the attributes C,C Sharp,D,D Sharp,...,B.

7) Spectrum Roll-off (RO) [17]: The spectral roll-off is defined as the frequency below which 85% of the magnitude distribution is concentrated. The roll-off is another measure of the spectral shape.

B. Recognition using Gaussian Mixture Model

In order to test the aforementioned features for texture recognition, a 3-component Gaussian Mixture Model (GMM) is used as the classifier in this paper. The GMM model is represented by a set of parameters $\lambda_m = \{\omega_i^m\ ,\ \mu_i^m\ ,\ \Sigma_i^m\}$, $i = 1, \ldots, M$, where ω_i^m, μ_i^m and Σ_i^m are the weight, mean and covariance of the i-th component, respectively. The diagonal covariance matrix can be referred to as the variance of the GMM model. These parameters can be estimated using Maximum a posteriori (MAP) estimation. For a group of textures S=$\{1,\ 2, ...,\ \text{S}\}$ represented by GMMs λ_1, λ_2, $\ldots$, λ_S, the objective of recognition is to find the model which has the maximum a posteriori probability for an incoming N-dimensional test data vector $\mathbf{X} = (x_1 \ldots x_N)^T$:

$$\hat{S} = arg \max_{1 \leq k \leq S} P_r(\lambda_k|\mathbf{X}) = arg \max_{1 \leq k \leq S} \frac{p(\mathbf{X}|\lambda_k)P_r(\lambda_k)}{p(\mathbf{X})} \quad (1)$$

where $p(\mathbf{X})$ is the same for all models. Assuming that $P_r(\lambda_k\ )$ is equal for each texture the recognition rule simplifies to:

$$\hat{S} = arg \max_{1 \leq k \leq S} P_r(\mathbf{X}|\lambda_k) \quad (2)$$

Using logarithms and the independence between observations, the recognition system result is computed as

$$\hat{S} = arg \max_{1 \leq k \leq S} \sum_{t=1}^{N} log(p(x_t|\lambda_k)) \quad (3)$$

C. Experimental Results

To evaluate the audio features, the first four seconds of all freehand queries are removed since the features should be applied only on the texture stroking signals. The remaining data points are used in a tenfold cross validation, where one query is used as test data and the other nine as training data. The same procedure is then repeated with the other nine combinations of training and test data. Each feature is tested individually using the aforementioned GMM model. Table I summarizes the achieved classification results, with accuracy denoting the overall fraction of correctly classified textures in the cross validation.

TABLE I. AUDIO FEATURE RECOGNITION RESULTS

Audio Feature	Dimensions	Accuracy [%]
MFCC	15	80.23
Rasta PLP	17	58.19
Spectrum Centroid	1	17.57
Spectrum Flatness	1	16.89
Chroma	12	8.49
LPCC	12	5.74
Spectrum Roll-Off	1	5.28

MFCCs yield the best results of all tested audio features with a total accuracy of 80.23%. They seem to be quite robust against the variance in our recorded free hand texture signals due to the uncontrolled scan parameters. Surprisingly, the LPC coefficients achieve only a very poor result of 5.74%, although they also represent the spectral envelope. The reason for this performance difference is not obvious and necessitates further investigation in future work. The poor result for Chroma seems to be explained by its strong audio-specific definition. The Spectrum Centroid and the Spectrum Flatness yield a considerable accuracy regarding the fact that it is only a one-dimensional feature. There are 43 different classes in our current evaluation and, hence, classification based on only a one-dimensional feature may lead to high confusion. Nevertheless, low-dimensional independent features might be still useful, if they are combined properly.

V. Conclusion

We introduce a haptic texture database containing controlled and uncontrolled acceleration recordings for 43 different texture materials using a tool-mediated interface. The purpose of this dataset is to analyze different feature candidates for texture recognition and retrieval systems where the human user has full control over the exploration movement. Our long-term goal is to identify features that are invariant against scan parameters and can be used without additional and costly sensors. The introduced database, thus, serves as an important basis for further investigations to achieve this objective.

As a preliminary feature analysis, we selected six well-established features from audio and speech recognition and use them in a texture recognition system based on GMMs. The experiments reveal that MFCC works best among the tested features with an achieved recognition accuracy of 80.2%. The presented results, however, are only preliminary as we consider only GMMs and a selected subset of available audio/speech recognition features in this paper. For future work, we plan to extend the list of features and to investigate, if they can be optimized to better match the haptic domain. Furthermore, we are planning to offset them against the features used in related work in the area of robotic surface recognition and also to compare machine learning algorithms other than GMM. The combination of different features, capturing different signal properties will obviously also help to improve the recognition accuracy.

Acknowledgment

This work has been supported, in part, by the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement no. 258941 and, in part, by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2013R1A1A2007601).

References

[1] S. Lederman and R. Klatzky, "Haptic perception: A tutorial," *Attention, Perception, & Psychophysics*, vol. 71, no. 7, pp. 1439–1459, October 2009.

[2] E. Saddik, "The Potential of Haptics Technologies," *IEEE Instrumentation & Measurement Magazine*, vol. 10, no. 1, pp. 10–17, April 2007.

[3] W. McMahan, J. M. Romano, A. M. A. Rahuman, and K. J. Kuchenbecker, "High frequency acceleration feedback significantly increases the realism of haptically rendered textured surfaces," in *IEEE Haptics Symposium (HAPTICS)*, March 2010, pp. 141–148.

[4] K. Kuchenbecker, J. Romano, and W. McMahan, "Haptography : Capturing and Recreating the Rich Feel of Real Surfaces," in *The 14th International Symposium on Robotics Research (ISRR)*, 2009.

[5] R. Chaudhari, B. Cizmeci, K. J. Kuchenbecker, S. Choi, and E. Steinbach, "Low Bitrate Source-filter Model Based Compression of Vibrotactile Texture Signals in Haptic Teleoperation," in *ACM Multimedia 2012*, 2012, pp. 409–418.

[6] J. Sinapov and V. Sukhoy, "Vibrotactile recognition and categorization of surfaces by a humanoid robot," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 488–497, 2011.

[7] N. Jamali and C. Sammut, "Majority voting: material classification by tactile sensing using surface texture," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 508–521, 2011.

[8] J. A. Fishel and G. E. Loeb, "Bayesian exploration for intelligent identification of textures," *Frontiers in neurorobotics*, vol. 6, no. 4, pp. 1–20, June 2012.

[9] J. M. Romano and K. J. Kuchenbecker, "Methods for Robotic Tool-Mediated Haptic Surface Recognition," in *IEEE Haptics Symposium (HAPTICS)*, 2014, pp. 49–56.

[10] H. Culbertson, J. J. L. Delgado, and K. J. Kuchenbecker, "One hundred data-driven haptic texture models and open-source methods for rendering on 3d objects," in *IEEE Haptics Symposium (HAPTICS)*, February 2014, pp. 319–325.

[11] M. Strese, C. Schuwerk, R. Chaudhari, and E. Steinbach. (2014) Haptic texture database. [Online]. Available: http://www.lmt.ei.tum.de/texture/

[12] J. M. Romano and K. J. Kuchenbecker, "Creating realistic virtual textures from contact acceleration data," *IEEE Transactions on Haptics*, vol. 5, no. 2, pp. 109–119, April 2012.

[13] H. Culbertson, J. M. Romano, P. Castillo, M. Mintz, and K. J. Kuchenbecker, "Refined methods for creating realistic haptic virtual textures from tool-mediated contact acceleration data," in *IEEE Haptics Symposium (HAPTICS)*, March 2012, pp. 385–391.

[14] C. G. McDonald and K. J. Kuchenbecker, "Dynamic simulation of tool-mediated texture interaction," in *World Haptics Conference (WHC)*, April 2013, pp. 307–312.

[15] N. Landin, J. M. Romano, W. Mcmahan, and K. J. Kuchenbecker, "Dimensional Reduction of High-Frequency Accelerations for Haptic Rendering," in *Proceedings of the Eurohaptics*, 2012, pp. 79–86.

[16] S. J. Lederman, "Hand movements: A window into haptic object recognition," *Cognitive Psychology*, vol. 19, no. 3, pp. 342–368, 2003.

[17] L. Lu, H. Jiang, and H. Zhang, "A robust audio classification and segmentation method," in *Proceedings of the ninth ACM international conference on Multimedia*. ACM, 2001, pp. 203–211.

[18] H.-G. Kim, N. Moreau, and T. Sikora, *MPEG-7 audio and beyond: Audio content indexing and retrieval*. John Wiley & Sons, 2006.

[19] D. A. Reynolds, "Speaker identification and verification using gaussian mixture speaker models," *Speech communication*, vol. 17, no. 1, pp. 91–108, 1995.

[20] L. R. Rabiner and B.-H. Juang, *Fundamentals of speech recognition*. PTR Prentice Hall Englewood Cliffs, 1993, vol. 14.

[21] V. Mitra and C.-J. Wang, "A neural network based audio content classification," in *International Joint Conference on Neural Networks*. IEEE, 2007, pp. 1494–1499.

[22] G. Guo and S. Z. Li, "Content-based audio classification and retrieval by support vector machines," *IEEE Transactions on Neural Networks*, vol. 14, no. 1, pp. 209–215, 2003.

[23] E. Wong and S. Sridharan, "Comparison of linear prediction cepstrum coefficients and mel-frequency cepstrum coefficients for language identification," in *Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing*. IEEE, 2001, pp. 95–98.

[24] H. Hermansky and N. Morgan, "Rasta processing of speech," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 4, pp. 578–589, 1994.

[25] M. Müller and S. Ewert, "Chroma toolbox: Matlab implementations for extracting variants of chroma-based audio features." in *ISMIR*, 2011, pp. 215–220.