Author(s): Tsatsishvili, Valeri; Burunat, Iballa; Cong, Fengyu; Toiviainen, Petri; Alluri, Vinoo; Ristaniemi, Tapani

Title: On application of kernel PCA for generating stimulus features for fMRI during continuous music listening

Year: 2018

Version:

Accepted Manuscript

Please cite this article as: Tsatsishvili Valeri, Burunat Iballa, Cong Fengyu, Toiviainen Petri, Alluri Vinoo, Ristaniemi Tapani.On application of Kernel PCA for Generating Stimulus Features for fMRI during Continuous Music Listening.*Journal of Neuroscience Methods* https://doi.org/10.1016/j.jneumeth.2018.03.014

On application of Kernel PCA for Generating Stimulus Features for fMRI during Continuous Music Listening

Research Article

Valeri Tsatsishvili[a], Iballa Burunat[b], Fengyu Cong[c,a], Petri Toiviainen[b], Vinoo Alluri[d], Tapani Ristaniemi[a]

[a]*Faculty of Information Technology, University of Jyvaskyla, Finland*

[b]*Department of Music, Art and Culture Studies, University of Jyvaskyla, Finland*

[c]*Department of Biomedical Engineering, Dalian University of Technology, China*

[d]*International institute of information technology, Hyderabad, India*

Highlights

- Nonlinear method for generating high-level musical features is proposed
- Novel features capturing rhythmic complexity and event synchronicity were identified
- Neural correlates of the new features have not been studied widely
- Nonlinear methods can find interesting stimulus features hidden from linear methods

Abstract

Background: There has been growing interest towards naturalistic neuroimaging experiments, which deepen our understanding of how human brain processes and integrates incoming streams of multifaceted sensory information, as commonly occurs in real world. Music is a good example of such complex continuous phenomenon. In a few recent fMRI studies examining neural correlates of music in continuous listening settings, multiple perceptual attributes of music stimulus were represented by a set of high-level features, produced as the linear combination of the acoustic descriptors computationally extracted from the stimulus audio.

New method: fMRI data from naturalistic music listening experiment were employed here. Kernel principal component analysis (KPCA) was applied to acoustic descriptors extracted from the stimulus audio to generate a set of nonlinear stimulus features. Subsequently, perceptual and neural correlates of the generated high-level features were examined.

Results: The generated features captured musical percepts that were hidden from the linear PCA features, namely Rhythmic Complexity and Event Synchronicity. Neural correlates of the new features revealed activations associated to processing of complex rhythms, including auditory, motor, and frontal areas.

Comparison with existing method: Results were compared with the findings in the previously published study, which analyzed the same fMRI data but applied linear PCA for generating stimulus features. To enable comparison of the results, methodology for finding stimulus-driven functional maps was adopted from the previous study.

Conclusions: Exploiting nonlinear relationships among acoustic descriptors can lead to the novel high-level stimulus features, which can in turn reveal new brain structures involved in music processing.

Keywords: functional magnetic resonance imaging (fMRI), naturalistic fMRI, kernel PCA, music stimulus, musical features, feature generation

1. Introduction

Music, as an important part of human experience, has long been of interest for neuroimaging community. Traditionally, temporal structure of stimulus presentation has been controlled and stimuli have been simplified in fMRI experiments in order to maintain control over independent variables. Such experimental design limits replicating experience of music and other multifaceted phenomena as they occur in real life and therefore, do not provide an optimal framework for studying sensory or cognitive processing mechanisms of continuous information flow. Consequently, interest towards naturalistic neuroimaging experiments, where participants are exposed to continuous complex stimuli is growing (e.g. see [1]-[3]). Music possesses several inter-related perceptual attributes including loudness, timbre, tonality, and rhythm. Many excellent works have examined neural correlates of each attribute in isolation to understand neural processing of specific musical features [4]-[6]. However, musical dimensions are rarely processed independently from each other and growing evidence suggests that the brain areas involved in music processing is not a simple integration over the networks associated to its isolated attributes [7], [8]. Therefore, segregation of neural correlates of the musical features should be achieved in natural music listening contexts in order to observe simultaneous processing of multiple features as it typically takes place in real life [8].

Functional MRI studies investigating activations in naturalistic music listening settings are scarce. Nevertheless, few attempts have been made recently. Abrams et al. [7] examined regions involved in music processing by comparing of synchronization between non-musician participants' brain responses while listening to the recordings of classical music (musical stimuli) as well as their phase scrambled and spectrally rotated versions (pseudo-musical stimuli). Results revealed that synchrony across subjects in cortical and subcortical regions was significantly greater in response to musical compared to the pseudo-musical stimuli. With different objectives, Alluri and colleagues in [8] studied neural correlates of multiple musical features in an experiment where 11 musicians listened to a piece of modern tango in the fMRI scanner. A set of 25 descriptors were computationally extracted from the music stimulus. The descriptors were subsequently summarized in nine principal components, or stimulus features, capturing high-level timbral, tonal, and rhythmic aspects of music. Perceptual relevance of the components was validated in the perceptual experiment. Neural correlates of the stimulus features revealed large-scale consistent foci of activations in cortical and sub-cortical areas. Analogous stimulus feature generation scheme was employed in several subsequent studies with similar experimental setup [9]-[11].

Interestingly, in [8] several brain areas showing high inter-subject consistency that failed to correlate with any of the stimulus features were also reported. Authors suggested an incomplete representation of stimulus as one possible explanation. In addition, two of the nine stimulus features had to be excluded in early stages of the analysis as they failed to significantly correlate with perceptual ratings from the validation experiment. These limitations might, among other possible reasons, stem from the existence of nonlinear patterns between the initially extracted 25 descriptors, which cannot be captured by linear PCA. In other domains, exploiting nonlinear patterns using KPCA-based features has shown to improve classification accuracy when compared to the performance of linear features [12]-[14]. This motivated us to explore application of nonlinear kernel PCA for generating perceptually relevant high-level features. This study employed the data from [8] (in the following will

2

be referred as original study) to generate alternative stimulus representation. Perceptual and neural correlates of the new stimulus features were obtained with the same methodology as in the original study, which enabled direct comparison of results.

Section 2 details methodology for generating nonlinear stimulus features as well as finding perceptual and neural correlates of the generated features. Section 3 presents behavioral correlates and stimulus-driven spatial maps. Section 4 highlights findings and limitation of the present study.

2. Materials and Methods

2.1. Acoustic Feature Extraction and preprocessing

Overall, 25 descriptors representing timbral, tonal, and rhythmic information were extracted from the overlapping windows. Window length was varied to facilitate different time scales of processing needed for different musical features. The shorter window length of 25ms with 50% overlap was selected for 20 low-level descriptors capturing polyphonic timbre of music. The longer windows spanning three seconds with 67% overlap were employed for the remaining more complex features capturing higher-level concepts, such as tonality and rhythm. These descriptors were adopted from the original study [8] where the complete list and descriptions are provided. All 25 descriptors in the initial set were centered and normalized with respect to their standard deviation, and downsampled to 0.5 Hz to match fMRI sampling rate. After the initial preprocessing, the features were subjected to KPCA. The generated kernel principal components were finally convolved with the canonical hemodynamic response function.

2.2. Kernel PCA

Kernel PCA is an extension of the linear PCA for nonlinear data distributions where mapping into linear subspace is not useful [15], [16]. The method has gained popularity as a tool for finding useful image representations with wide variety of applications such as edge detection, handwriting recognition, and classification of various medical images.

The most straightforward way to do nonlinear extension of PCA is to introduce a nonlinear mapping to (generally) higher dimensional feature space: $X \rightarrow \phi(x)$, calculate covariance $V = \phi(x)\phi(x)^T$ in the feature space and then solve the following eigenvalue problem:

$$Cv = \lambda v \tag{1}$$

Such mapping can become very expensive or infeasible as the dimensionality of X increases. However, in certain cases it is possible to bypass the mapping step and directly calculate the dot product in the feature space. This is achieved by introducing a kernel function that replaces dot product in the feature space:

$$K(x,x) = E\{\phi(x)\phi(x)^T\} = \frac{1}{m}\sum_{i=1}^{m}\phi(x_i)\phi(x_i)^T \tag{2}$$

This procedure, called 'Kernel trick' reduces mapping and dot product operations to kernel function estimation and can be applied as nonlinear extension of any linear method that depends exclusively dot products. Once kernel function is estimated, inserting (2) in (1) reformulates the eigenvalue problem:

$$\frac{1}{m}\sum_{i=1}^{m}\phi(x_i)\phi(x_i)^T v = \lambda v \tag{3}$$

By considering that all solutions of v in (3) lie in the span of $\phi(x)$:

$$v = \sum_{i=1}^{m}a_i\phi(x_i) \tag{4}$$

3

and substituting (4) into (3), the eigenvalue problem can be expressed as:

$$Ka = \lambda a \tag{5}$$

Finally, principal components can be obtained by projecting the sample onto the eigenvectors of the covariance matrix in the feature space:

$$\langle v, \phi(x) \rangle = \sum_{i=1}^{m} a_i K(x_i, x) \tag{6}$$

Many nonlinear functions are suitable as the kernel functions. The most common, Gaussian kernel was selected here:

$$k_{ij} = \exp(-\frac{\|x_i - x_j\|^2}{2\varepsilon^2}) \tag{7}$$

where k_{ij} are the elements of the kernel matrix, and x_i, x_j are d dimensional instances of X. The ε parameter was set as median of the minimum value of the distances between the data points i.e. $\varepsilon = \underset{i}{\text{median}} \underset{i \neq j}{\min} \|x_i - x_j\|$.

Consequently, 10 kernel PCA features were selected for further analysis, retaining about 90% of the variance in the data. More thorough derivation of kernel PCA can be found in e.g. [14].

2.3. Finding Perceptual correlates of KPCA features

Perceptual interpretation of the generated kernel PCA features is not straightforward. As mentioned above, in the original study Alluri et al. applied PCA to the 25 initially extracted descriptors. The relevant musical percepts corresponding to each produced principal component (PC) was estimated by observing principal component loadings, i.e. contribution of each of the 25 initial features to the given PC. Consequently, the nine PC-s were associated with the following musical percepts: Fullness, Brightness, Timbral Complexity, Activity, Rhythmic Complexity, Event Synchronicity, Pulse Clarity, Key Clarity, and Mode. First four can be loosely defined as four dimensions of polyphonic timbre, following three represent different aspects of musical rhythm, and the remaining two relates to the tonality of music. The applied perceptual labels except Mode (excluded early from the analyses) were validated in the subsequent listening experiment, where 21 musicians rated short musical excerpts representing varying levels of each percept on scales from 1 to 9. To select the musical excerpts, entire stimulus piece was divided by six second moving windows with one-second step size, producing over 470 windows. Next, average PC values from all windows were extracted, sorted and 30 values were equidistantly sampled from the entire range. Principal components labelled as Rhythmic Complexity and Event Synchronicity failed to correlate with corresponding perceptual ratings and were excluded, leaving six validated features that were used in subsequent analyses of the original study.

Unlike PCA, KPCA operates in higher dimensional feature space without explicitly mapping the input data in it. Thus, interpreting nonlinear acoustic components by observing the weights of the initial 25 descriptors in the resulting kernel PC-s is not applicable here. Following a more heuristic approach, we employed the 9 sets of 30 short stimulus excerpts and human ratings from the listening experiment conducted in the original study. To this end, average values of the 10 kernel PCs were extracted from the musical excerpts and correlated to the corresponding perceptual ratings. Statistical significance of correlations was estimated by permutation tests: for each feature, average feature values from randomly sampled 30 windows were correlated with randomly selected rating multiple (100 000) times. The threshold of significance for correlation coefficients was derived from the resulting distribution.

In addition to the significance of correlations, we required for a given set of 30 samples to capture majority (at least 80%) of the entire range of feature values. Otherwise, correlation (even significant) was rejected. This constraint was introduced to be concordant with the sampling criterion originally applied to the linear features. As mentioned above, Alluri et al. [8] selected stimulus excerpts such that the entire dynamic range of the PCA features was equidistantly sampled, and consequently participants of the experiment rated the excerpts spanning all levels of the given feature value. The same set of audio excerpts represent arbitrary sampling for KPCA features, which in turn has two limitations if directly correlated with human ratings: 1) it increases chance of spurious correlation and 2) by possibly capturing a fraction of the feature range, it violates existing systematic relationship between the ratings and corresponding acoustic components.

2.4. Finding stimulus related brain activations

The dataset analyzed in this study was collected and preprocessed in the original study and consists of fMRI scans of eleven healthy musicians (mean age: 23.2; SD: 3.7; 5 females). Participants were instructed to listen attentively to an 8.5-minutes-long instrumental composition by Astor Piazzolla and maintain their gaze on the screen. The fMRI measurements were made in 3T scanner at sampling frequency of 0.5 Hz. Obtained fMRI scans went through the preprocessing routine, which included realignment, spatial normalization, smoothing and high-pass filtering at cutoff frequency of 0.008Hz to reduce effects of scanner drift.

Following the approach employed in the original study, stimulus-driven activation maps were first obtained per subject and per stimulus feature by correlating fMRI voxel time series to the KPCA features. Correlation coefficients were converted to z score maps using Fisher's z transformation, which were adjusted for serial correlations and converted to p-maps. Individual subject maps were then pooled to group maps using Fisher's p-value technique, such that one functional map was associated to each KPCA feature. The group maps were thresholded at p<0.001. Finally, significance thresholds for group maps were adjusted for multiple comparisons using cluster size thresholding. More detailed description of the analysis pipeline is provided in [8].

The obtained group-level maps driven by the KPCA features (K-maps) were compared with the PCA-driven maps (PC-maps) reported in the original study. Similarity between the spatial maps was evaluated at global and individual level. At global level, a thresholded individual K-maps and PC-maps were first combined, and the overlap between the combined maps was estimated by:

$$O_{kp} = \frac{K \cap P}{|K|} \qquad (8)$$

where K and P represent the voxels showing above-threshold activity in K-maps and PC-maps respectively, $|K|$ is the number of active voxels in K-map. Such global comparison provided an overview of the shared and exclusive regions of activity. For the individual analysis, spatial correlation was calculated between the continuous valued (unthresholded) maps. Significance of correlations was assessed by permutation tests. The procedure was similar to the one implemented in [11]: one feature from each feature set (PCA and KPCA) was randomly selected and phase scrambled. Each randomized feature was voxel-wise correlated to the fMRI data from all participants. The subject-specific maps were subsequently averaged, producing one group map per feature. Finally, correlation between the pair of group maps was calculated. The process was repeated 11 000 times and the significance threshold for correlation was estimated from the resulting distribution.

3. Results

3.1. Perceptual Correlates

Table 1 contains correlations between the nonlinear features extracted from the audio excerpts and corresponding perceptual ratings. As evident from the table, majority of KPCA features were correlated significantly to the ratings of at least one musical percept. However, the constraint on spanned dynamic range was not satisfied in all cases and accordingly, K_4, K_5, and K_{10} were eliminated from further analysis. From the remaining features, K_1 seemingly captured all four dimensions of polyphonic timbre and Pulse Clarity; K_2 correlated with Brightness and Event Synchronicity; K_3 and K_6 were associated with the perceived Pulse Clarity and Key Clarity respectively, whereas K_9 was correlated to Brightness and Rhythmic Complexity. Finally, no significant correlations were observed for K_7 and K_8, which might indicate that they relate to the perceptual attributes that were not rated in the perceptual experiment.

	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8	K_9	K_{10}
Fullness	0.78**	-0.23	0.33	0.21	-0.14	-0.27	0.09	-0.2	-0.09	0.13
Brightness	0.39**	0.48**	-0.02	0.1	-0.47	0.02	-0.03	0.33	0.39**	-0.17
T. Complexity	0.73**	0.08	0.22	0.33	-0.05	-0.07	-0.02	-0.22	0.11	-0.28
R. Complexity	-0.1	0.32	0.12	0.1	-0.28	-0.35	-0.33	0.31	0.38**	-0.03
Key Clarity	0.31	0.04	0.2	-0.17	-0.1	-0.53**	0.29	-0.15	-0.22	-0.21
Pulse Clarity	0.74**	-0.14	-0.38**	0.27	0.28	0.03	0.01	-0.04	-0.2	0.29
E. Synchronicity	0.16	-0.65**	0.21	0.13	0.38	0.21	0.11	-0.32	-0.32	-0.38
Activity	0.8**	-0.11	0.23	0.44	-0.01	-0.1	-0.08	0.08	0.28	-0.3

*Table 1. Correlations between KPCA and perceptual ratings. Significance threshold r=±0.37 (p<0.05). ** indicates cases where correlation is significant, and the additional constraint holds.*

Strong association of the kernel features with the Event Synchronicity and Rhythmic Complexity ratings is very interesting finding, since these were among the musical percepts that did not validate in the perceptual experiment in the original study (see Table 2 in [8]). Neural correlates of these features are introduced in the following section.

3.2. Neural correlates

Similarity between the spatial maps at the global level is summarized in Table 2, where the sizes of combined binary maps and the overlap is provided. The size of the spatial map is defined by the number of voxels above the threshold of activation. Therefore, as the table shows, 76% of 9863 active voxels in the combined K-map are shared with the combined PC-map. Figure 1 depicts the shared and exclusive regions of the two combined activation maps. As can be seen in the figure, auditory areas occupied the largest portion of the common areas. Elsewhere, overlapping clusters were observed in cerebellum, left precentral and bilateral postcentral gyri, bilateral inferior parietal gyrus, and right putamen. The unique areas associated with KPCA features (upper row in Figure 1) were mostly concentrated in frontal lobe, including bilateral superior frontal gyrus (SFG), middle frontal gyrus (MFG), inferior frontal gyrus (IFG), right precentral gyrus, as well as in cerebellum (crus I and II, vermis III, VII and VIII). In addition, activated clusters in visual processing-related areas (precuneus, right lingual gyrus and left middle and superior occipital gyri), bilateral anterior cingulate and paracingulate gyrus, and right paracentral lobule were also present.

Figure 2 depicts activation maps corresponding to features K_2 (Rhythmic Complexity) and K_9 (Event Synchronicity), whereas correlations between K-maps and PC-maps are provided in Table 3. These musical percepts were not represented by PCA features in the original study as corresponding principal components failed correlate significantly with the perceptual ratings. Interestingly, both features also predicted perceived brightness of timbre. Functional map of K_2 consists of activations in auditory cortices (in the left hemispheric prominently), visual-

processing related areas (precuneus) and some cerebellar areas. Left SMA and the right IFG tend to deactivate in response to this feature. As shown in Table 3, no significant correlations were observed between K_2-map and any of the PC-maps, although thresholded map shared predominantly auditory areas with timbral maps. In addition, K_2 map featured regions in right medial inferior orbitofrontal gyrus (negative correlation) and the left middle and superior occipital gyri (positive correlation), not reported in the original study. Similarly, thresholded K_9-map featured some overlap with timbral maps in auditory regions. Bilateral STG, MTG, right HG, and right superior temporal pole were observed to activate in response to K_9, while small areas in left inferior parietal supramarginal and angular gyri, right postcentral gyrus, and right supramarginal gyrus showed deactivation.

From the remaining features, the largest activation map was driven by K_1. Large scale activations included areas bilateral STG, Heschl's gyrus (HG), rolandic operculum (RO), supramarginal gyrus, superior temporal pole, insula, MTG as well as some cerebellar regions. These areas were associated to processing of polyphonic timbre in the original study. Indeed, Table 3 reveals high similarity of with the activation maps associated to timbral features and Pulse clarity. Negatively correlated regions to K_1 were located mainly in frontal and parietal lobes. We also observed few active clusters in occipital gyrus and orbitofrontal cortex that were not reported in the previous study.

Functional map of K_3 consisted of only a few hundred voxels, predominantly showing auditory area (bilateral STG and left MTG) and small cluster in postcentral gyrus.

Finally, feature K_6 was negatively correlated with perceived Key Clarity. Increased activation in clusters located in the left precuneus, left angular gyrus, left supramarginal gyrus and left lateralized deactivations in cerebellum, inferior frontal gyrus, median cingulate and paracingulate gyrus were observed in response to this feature. In the right hemisphere, deactivations in only two small clusters in right insula and right putamen were found.

4. Discussion

The present study explored application of kernel PCA for finding new perceptually meaningful auditory stimulus features for naturalistic fMRI experiment. In the previously published study [8] fMRI data from music listening experiment were collected to examine perceptual and neural correlates of the high-level musical features. These features were generated as the linear combination of 25 acoustic descriptors extracted from the music stimulus. In the present study, the same fMRI data was employed, and a new set of stimulus features were generated as nonlinear combinations of the initial acoustic descriptors. Perceptual and neural correlates of the new stimulus features were compared with the results in the original study. While the present study was built upon the data and methodology of the original study, exact replication of the previously reported spatial maps were not expected since stimulus features in the two studies were different.

First, perceptual correlates of the stimulus features were examined. We found significant correlations between the majority of kernel PC-s and the perceptual ratings. More interestingly, the ratings of Event Synchronicity and Rhythmic Complexity were explained better by the kernel PC-s than by the linear PC-s. This finding indicates that the KPCA features are capable of capturing high-level musical percepts hidden from the linear PCA counterparts. Interestingly, it was also observed that KPCA features were significantly correlated with multiple perceptual ratings (see Table 1). While this can be beneficial for compactness of representation, it might also make features difficult to interpret. For example, as observed above K_1 feature was simultaneously associated to perceived timbre and rhythm, which is not straightforward to explain. Nonetheless, existence of a latent link between perceived clarity of the main beat and color of the timbre is possible. Intriguingly, multiple associations were also found when linear acoustic components were correlated to all perceptual ratings. For example, in the original study acoustic component labeled as Fullness was reported to be significantly correlated with Fullness rating ($r = 0.80; p < 0.001$). We found that Fullness component can successfully predict Timbral Complexity, Pulse Clarity and Activity ratings ($r = 0.68, \ p < 0.001; \ r = 0.74, \ p < 0.001; r = 0.79, \ p < 0.001$). Lack of data is the most likely

explanation of such results. As mentioned in Section 2.3, in the perceptual experiment only 30 short audio samples were rated per scale. For the long stimulus piece lasting over seven minutes, 30 samples might capture large-scale dynamics of acoustic features but will miss fine details. This coarse representation is then further smoothed when the ratings are averaged across participants. Consequently, the ratings of different musical percepts will be more likely to exhibit very similar dynamics capturing only the most significant changes in music stimulus. Therefore, while served well for validation of already applied labels in the original study, the available behavioral rating data are probably not optimal to use for our purposes. The solution would be to collect more ratings in a new perceptual experiment, where the new sets of representative music excerpts are selected by equidistantly sampling KPCA features. Furthermore, the obtained excerpts should be rated on more perceptual scales than was available here, and more rigorous testing on inter-subject agreement should be performed before averaging the ratings. Alternatively, averaging the ratings can be avoided and instead correlation between a given feature and individual ratings can be averaged.

Another important issue related to the feature generation is the selection of the kernel function and its parameters. Due to the absence of the evaluation strategy that would quantify 'interestingness' of the features in early stages of this study, the most commonly applied Gaussian kernel was chosen as a starting point. Kernel width for the Gaussian kernel was also estimated based on the commonly applied heuristics. There is a plenty of room for improvements in this regard. For instance, both the kernel and parameter selection can be defined as an optimization problem where objective is to maximize correlations between the generated features and the behavioral data. This should also reduce the number of discarded features that fail to correlate with the perceptual ratings.

	Global map size (voxels)	Overlap
KPCA	9863	0.76
PCA	11468	

Table 2. Overlap of global thresholded KPCA map with the global thresholded PCA map.

	K_1	K_2	K_3	K_6	K_9
Fullness	0.98*	-0.19	0.2	-0.09	0.42
Brightness	0.9*	0.41	0.28	-0.12	0.46
T. Complexity	-0.69*	-0.24	-0.8**	-0.16	-0.67**
Key clarity	-0.04	-0.03	0.02	-0.41	-0.17
Pulse clarity	0.68*	-0.16	0.27	0.45	0.33
Activity	0.99*	-0.13	0.12	-0.12	0.44

Table 3. Correlations between continuous K-maps and PC-maps. * denotes significant correlations (thresholds: r=±0.47 at p<0.05; r=±0.63 at p<0.01).
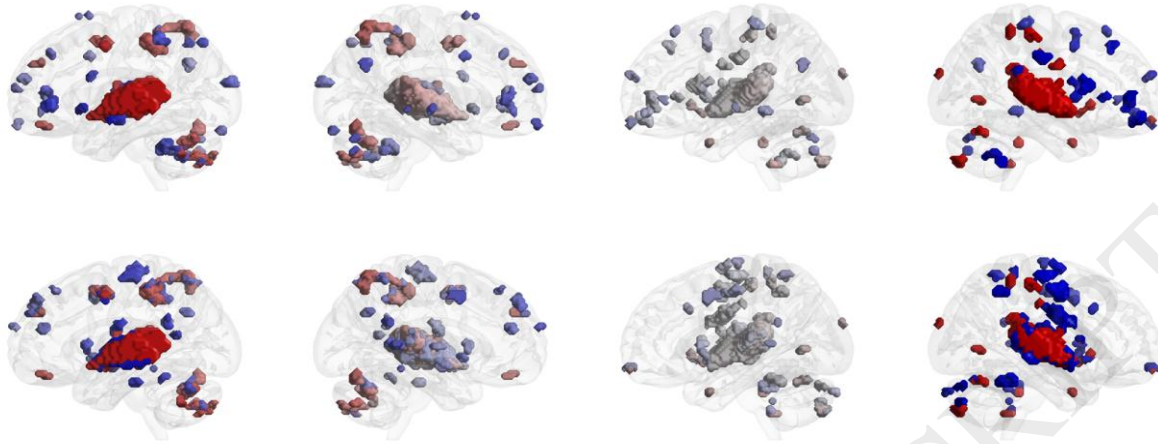
Figure 1. Shared (red) and exclusive (blue) regions between combined K-map and combined PC-map. Exclusive regions in combined K-map and PC-map are shown on top and bottom row, respectively. Note that polarity information is discarded here.
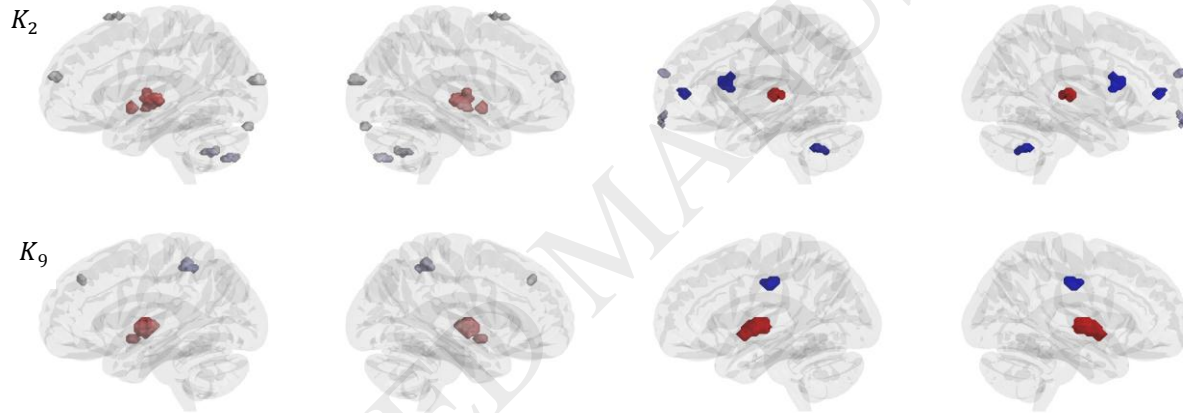

Figure 2. Spatial maps of the selected KPCA features. Note that polarity information is preserved here. Red color denotes positive correlations and blue – negative correlations.

As mentioned above, K_2 and K_9 predicted Event Synchronicity and, Rhythmic Complexity ratings better than the linear PCA features. Both are high-level musical percepts characterized by rhythmic fluctuations of energy in different frequency bands (fluctuation spectrum). Event synchronicity is the centroid of the fluctuation spectrum and roughly relates to the average of periodicities in different bands, while Rhythmic Complexity relates to the noisiness of the fluctuation spectrum. In addition, both kernel features were correlated with Brightness ratings. As these features were absent in the original study, neural correlates cannot be compared with the previous findings. Nevertheless, auditory areas comprise the majority of positive correlates for both feature maps. From the regions associated with K_2, previous studies on neural processing of rhythm implicate auditory cortex, SMA and cerebellum to be involved in rhythm perception [17]-[20] . Recruitment of frontal areas such as the right IFG may denote aspects of temporal attention processing during the perception of complex rhythms. Work by Chen et al. [20] revealed that frontocortical areas including the IFG were modulated by complexity of the rhythm during an auditory-motor synchronization task, and that they were more extensively recruited in musicians than in non-musicians. They argued this finding to underlie a greater involvement of working memory resulting from musicians' superior ability to organize temporal structure. Interesting to highlight is also activation of visual processing regions (left middle and superior occipital gyri) in response of increased K_2 and deactivation of right

9

supramarginal gyrus in response to K_9, since visual areas have not been reported active in response to rhythm processing thus far. Nevertheless, the precuneus, another area associated with visuo-spatial processing and attention, was also found to be negatively correlated with the clarity of the pulse in the original study. Given the lack of studies on rhythm complexity, and above-mentioned issues on reliability of perceptual interpretation, further investigation is needed in order to achieve a proper interpretation behind the functional neuroanatomy of these musical percepts.

To conclude, this study demonstrated that kernel PCA has the potential to produce features that capture novel musical percepts hidden from the linear features. Namely, two new rhythmic features representing periodicities in different bands of the frequency spectrum were identified. In addition to previously known anatomical regions involved in rhythm processing, visual processing areas were found to be associated with the new features. However, we recommend the highlighted limitations of this study to be addressed in future investigations before generalizing findings.

Acknowledgements

References

[1] U. Hasson *et al*, "Intersubject synchronization of cortical activity during natural vision," *Science,* vol. 303, *(5664),* pp. 1634-1640, Mar 12, 2004.

[2] H. J. Spiers and E. A. Maguire, "Decoding human brain activity during real-world experiences," *Trends Cogn. Sci. (Regul. Ed.),* vol. 11, *(8),* pp. 356-365, 2007.

[3] U. Hasson and C. J. Honey, "Future trends in Neuroimaging: Neural processes as expressed within real-life contexts," *Neuroimage,* vol. 62, *(2),* pp. 1272-1278, 2012.

[4] P. Janata *et al*, "The cortical topography of tonal structures underlying Western music," *Science,* vol. 298, *(5601),* pp. 2167-2170, Dec 13, 2002.

[5] J. D. Warren *et al*, "Separating pitch chroma and pitch height in the human brain," *Proc. Natl. Acad. Sci. U. S. A.,* vol. 100, *(17),* pp. 10038-10042, Aug 19, 2003.

[6] A. Caclin *et al*, "Separate neural processing of timbre dimensions in auditory sensory memory," *J. Cogn. Neurosci.,* vol. 18, *(12),* pp. 1959-1972, 2006.

[7] D. A. Abrams *et al*, "Inter-subject synchronization of brain responses during natural music listening," *Eur. J. Neurosci.,* vol. 37, *(9),* pp. 1458-1469, 2013.

[8] V. Alluri *et al*, "Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm," *Neuroimage,* vol. 59, *(4),* pp. 3677-3689, Feb 15, 2012.

[9] V. Alluri *et al*, "From Vivaldi to Beatles and back: Predicting lateralized brain responses to music," *Neuroimage,* vol. 83, pp. 627-636, 2013.

[10] P. Toiviainen *et al*, "Capturing the musical brain with Lasso: Dynamic decoding of musical features from fMRI data," *Neuroimage,* vol. 88, pp. 170-180, 2014.

[11] I. Burunat *et al*, "The reliability of continuous brain responses during naturalistic listening to music," *Neuroimage,* vol. 124, pp. 224-231, 2016.

[12] G. S. Sidhu *et al*, "Kernel Principal Component Analysis for dimensionality reduction in fMRI-based diagnosis of ADHD," *Frontiers in Systems Neuroscience,* vol. 6, 2012.

[13] T. Mu, A. K. Nandi and R. M. Rangayyan, "Classification of breast masses via nonlinear transformation of features based on a kernel matrix," *Med. Biol. Eng. Comput.,* vol. 45, *(8),* pp. 769-780, 2007.

[14] B. Scholkopf, A. Smola and K. Müller, "Kernel principal component analysis," in *Advances in Kernel Methods-Support Vector Learning,* 1999.

[15] K. Muller *et al*, "An introduction to kernel-based learning algorithms," *Neural Networks, IEEE Transactions On,* vol. 12, *(2),* pp. 181-201, 2001.

[16] B. Scholkopf *et al*, "Input space versus feature space in kernel-based methods," *Neural Networks, IEEE Transactions On,* vol. 10, *(5),* pp. 1000-1017, 1999.

[17] S. L. Bengtsson *et al*, "Listening to rhythms activates motor and premotor cortices," *Cortex,* vol. 45, *(1),* pp. 62-71, 2009.

[18] J. A. Grahn and J. B. Rowe, "Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception," *J. Neurosci.,* vol. 29, *(23),* pp. 7540-7548, Jun 10, 2009.

[19] J. A. Grahn and M. Brett, "Rhythm and beat perception in motor areas of the brain," *J. Cogn. Neurosci.,* vol. 19, *(5),* pp. 893-906, 2007.

[20] J. L. Chen, V. B. Penhune and R. J. Zatorre, "Listening to musical rhythms recruits motor regions of the brain," *Cereb. Cortex,* vol. 18, *(12),* pp. 2844-2854, Dec, 2008.