Perceptual Importance of Karhunen-Lòeve Transformed Multichannel Audio Signals

Lars Henning¹, Yu Jiao, Slawomir Zielinski and Francis Rumsey

Institute of Sound Recording, University of Surrey, Guildford, GU2 7XH, UK

lars@focusrite.com
y.jiao@surrey.ac.uk
s.zielinski@surrey.ac.uk
f.rumsey@surrey.ac.uk

ABSTRACT

The Karhunen-Lòeve Transform (KLT) can be used to reduce the interchannel redundancy of multichannel audio signals. For this paper, the perceptual importance of Karhunen-Lòeve transformed multichannel audio signals was systematically studied using two experiments. The first experiment investigated the perceptual effects caused by removing some KLT eigenchannels. The results showed that some eigenchannels are not perceptually important and consequently can be discarded with minimal degradation of basic audio quality. The second experiment involved further investigation on the perceptual effect of KLT processing on the audio quality of multichannel audio as a function of the nature of the multichannel audio and eigenvalue extraction methods of KLT processing. It was also attempted to establish the relationship between the order of perceptual importance and the order of statistical importance of KLT eigenchannels.

¹ Lars Henning is now employed by Focusrite Audio Engineering Ltd, High Wycombe, HP12 3FX, UK.
1. INTRODUCTION

Transmission of High Definition Television (HDTV) is limited by digital data rate, much like analogue television has always been limited by analogue transmission bandwidth. This is equally true for the audio that accompanies television picture. For this reason, it is clear that an efficient coding scheme is needed for the storage and transmission of multichannel audio. The requirements of this scheme include optimised methods of data reduction as well as advanced quality scalability.

In the past, optimised perceptual coding methods based on psychoacoustic models and on redundancy within channels (e.g., entropy coding) have been utilised to reduce the bit rate of multichannel audio [1]. However, in most commonly used technologies (e.g. Dolby AC-3 and AAC), only the redundancy between channel-pairs has been taken into account. Some experimental results have shown that high correlation exists between every pair of channels in all frequency regions in some kinds of recording [2]. Recent studies by Yang et al. have led to the development of a new algorithm called the Modified Advanced Audio Coding Karhunen Lòeve Transform (MAACKLT), which incorporates a modified AAC algorithm with offline Karhunen-Lòeve transformation in the preprocessing stage [3]. As KLT can transform a set of correlated signals into a set of uncorrelated (orthogonal) signals, MAACKLT can remove interchannel redundancy of all channels to improve the successive coding performance.

In addition to signal decorrelation, KLT has another very valuable property: the KL-transformed signals are ordered hierarchically according to the amount of signal variation they explain. The first KL-transformed signal is a linear combination of the original signals that explains the greatest amount of variance. The second KL-transformed signal defines the next largest variance. The number of KL-transformed signals is equal to, or less than, the number of original signals. The total amount of variance of KL-transformed signals equals that of the original signals.

From a statistical viewpoint, variance can be used as a measure of how much information a signal contains. Therefore, it is reasonable to claim that the KL-transformed signals are ordered according to their statistical importance. This hierarchy of statistical importance encourages the idea that when transmitting data comprised of a fixed number of audio signals, dropping the least important KL-transformed audio signals2 (those which are lower in the hierarchy) will enable retention of the most important information and will only cause minimal degradation of the audio quality. However, it cannot be directly deduced that the mathematical character of KLT will lead to the same effect on people's perception of the quality of multichannel audio. Unfortunately, little research has been carried out on the perceptual effect of KLT in the context of multichannel audio.

In this paper, the perceptual importance of Karhunen-Lòeve transformed multichannel audio signals was systematically studied using two experiments. As it is the de facto standard in the consumer audio market, the 5.1-channel setup (see Figure 1) was used as the representative multichannel audio format in these experiments. The LFE channel was omitted in this work.

Figure 1: Loudspeaker set-up according to the ITU-R BS. 775 Recommendation

The two experiments are presented independently in the following sections. Each can be considered a self-contained part, including respective conclusions and discussions. The last section of the paper gives a summary of the two experiments.

2 The KL-transformed signals in the context of multichannel audio are referred to as eigenchannels throughout this paper.
2. EXPERIMENT I

The first experiment involved an investigation of the perceptual importance of eigenchannels for KL-transformed multichannel audio with high interchannel correlation. The purpose was to investigate how KLT eigenchannel removal will affect listener perception of multichannel audio quality and fidelity. Beginning with the least statistically significant, eigenchannels were removed until only the most significant eigenchannel remained. Three perceptual quality factors were studied: Basic Audio Quality (BAQ), Frontal Spatial Fidelity (FSF) and Surround Spatial Fidelity (SSF). BAQ refers to any and all degradations of audio quality while FSF and SSF involve only perceived differences of spatial characteristics excluding timbral quality.

The research questions for Experiment I were:

1. How does KLT eigenchannel removal affect listener perception of basic audio quality, frontal spatial fidelity, and surround spatial fidelity for the selected multichannel audio?

2. Is there a significant difference between degradation of basic audio quality, frontal spatial fidelity, and surround spatial fidelity caused by KLT eigenchannel removal?

2.1. Selection of Audio Excerpts

Three multichannel audio excerpts were chosen for the listening test based upon two main criteria: interchannel correlation and spatial characteristics.

The first criterion was the amount of interchannel correlation, which may affect the performance of the KLT processing. If the original signals are highly correlated, the first few eigenchannels will represent most of the variance and the last eigenchannels will contain a very small amount of variance. It was expected that the audio quality degradation would be small if the last eigenchannels were removed for this type of recordings. However, if the original signals are totally uncorrelated, the KLT will produce eigenchannels identical to the original signals but in an order according to the amount of variance they contain. For these recordings, KLT is not useful for data-rate reduction, and audio degradation would be the same as it would be with the removal of normal audio signals.

The interchannel correlation can be measured using correlation matrices. The correlation coefficients in these matrices show the amount of correlation between any two audio channels for the duration of the window length (in this case, the window length was the length of the excerpt). Each correlation coefficient within a matrix falls between \([-1, 1]\) where a coefficient of ‘1’ represents 100% correlation, ‘0’ represents no correlation and ‘-1’ represents 100% correlation out of phase (essentially inverted). However, it is not easy to compare the total interchannel correlation of different excerpts using correlation matrices. Since the degree of interchannel correlation will affect the distribution of variance among the eigenchannels, the percentage of variance represented by the first eigenchannel can be used as a single-value measurement of the interchannel correlation.

Excerpts with high interchannel correlation were selected for this experiment. Scree plots are presented in Appendix A1, which show the percentage of variance explained by the each of the five eigenchannels. It can be seen that for each of the three excerpts, the 1\(^{st}\) eigenchannel represent a high percentage of the total variance and the 4\(^{th}\) and 5\(^{th}\) eigenchannels represent a low percentage of the total variance. The percentages of variance were calculated using only the covariance matrix-based eigenvalue extraction method\(^3\).

The second aim of this experiment was to investigate any possible correlation between the spatial character of each excerpt and the ability of listeners to detect the removal of certain eigenchannels. Therefore, the second criterion for selection of audio selection was that the excerpts should cover a variety of spatial characteristics common to surround sound programme material. The spatial characteristics of multichannel audio were described and classified using Rumsey’s Scene-Based Paradigm [5]. The most basic spatial scenes are labelled FB and FF. The FB scene describes the case where front channels reproduce “foreground” audio content (close and clearly perceived audio sources), whereas rear channels contain “background” audio content.

\(^3\) The eigenvalue extraction methods are discussed in section 3.2.
(reverberant sounds, not clear, “foggy”, quieter than the front ones). The FF scene describes a recording in which the listener is surrounded by clearly identifiable audio sources (foreground audio content both from front and rear directions) [9]. From previous research, it was found that some recordings had a predominant centre channel (voices or solo instruments). These kinds of recordings are perceived very differently from those without predominant centre channel content in terms of spatial character. Therefore, recordings with predominant centre channel were discriminated from basic FB and FF recording as FB+C and FF+C.

The three chosen excerpts exhibit the following surround programme types and spatial scenes:

Excerpt 1: Pop Music FF
Excerpt 2: Movie FB+C
Excerpt 3: Pop Music FB

The details of these excerpts are presented in a table in Appendix A1.

2.2. Processing of Audio Material

Once the three excerpts were selected, it was necessary to prepare seven ‘audio processes’ for use in the MUSHRA listening test. These processes included:

Reference – the unprocessed reference version
Process A – 5th eigenchannel removed
Process B – 4th & 5th eigenchannels removed
Process C – 3rd, 4th, & 5th eigenchannels removed
Process D – 2nd, 3rd, 4th, & 5th eigenchannels removed
Process E – Hidden Reference
Process F – Hidden Anchor

A simplified overview of this process is illustrated in Figure 2. The gain of each eigenchannel was either set to ‘0’ or ‘1’, were a ‘0’ effectively removes a particular eigenchannel, and ‘1’ sets an eigenchannel’s gain to unity. A KLT gain matrix was used to generate each of the four audio processes for each of the three chosen excerpts. After the gain matrix was applied to a transformed excerpt, an inverse KLT was applied in order to convert the excerpt back into 5 loudspeaker signals.

\[
k = \frac{\sum_{i=1}^{n} |\text{Process} - \text{Reference}|}{n}
\]  

(Eq. 1)

In Eq. 1 above, both Process and Reference refer to 5-channel matrices. A k-matrix was created for the 3 excerpts. These matrices were used to create the k-graphs presented in Appendix A3.

The mean difference amounts shown in the k-graphs allow the effects of the KLT on each audio channel to be quantified. They were used in order to predict the listeners’ ability to detect differences between the processes and the reference. It was predicted that the greater mean differences shown in Processes C and D would be more easily detected by the listeners.

2.4. Experimental Design

A Multiple Stimulus Hidden Anchor and Reference (MUSHRA) listening test was devised and administered in accordance with the standard ITU-R BS.1534-1 Recommendation [6].
2.4.1. Selection of Listening Subjects

A total of 23 normal hearing Tonmeister (Music & Sound Recording) students were selected to take part in the MUSHRA listening test. Experienced listeners were chosen, as the aim was to investigate subtle perceptual differences which are best assessed by experienced listeners. It was decided that a minimum of 15 listeners would be needed in order to provide a sample large enough for thorough data analysis.

2.4.2. The MUSHRA Graphical User Interface

The MUSHRA test was chosen because it allowed the listeners to make quick, repeated comparisons of the audio processes and the reference audio. After the requirements of the interface were established, a graphical user interface (GUI) was then developed. The MUSHRA interface for Experiment I is shown below in Figure 3.

![MUSHRA Interface](image)

Figure 3: MUSHRA interface for Experiment I.

2.4.3. The Listening Test Environment

The listening tests were conducted in Studio 3 of the PATS building at the University of Surrey, Guildford, UK. The listening position was optimised as specified by the BS.775-1 standard [7]. The listeners sat facing the computer monitor, on which the MUSHRA interface was displayed. Following the presentation of verbal and written instructions, a training phase was administered for each listener. Subjects were reminded that at least one excerpt in each section should be given a rating of ‘100’ and that ‘spatial fidelity’ was to exclude any differences in ‘timbral quality’. The playback level was set to an initial ‘preferred listening level’ on the first day of testing as recommended by the BS.1116-1 standard [8]. The five loudspeakers were level-aligned using broadband pink noise, reproduced through individual channels and measured in SPL at the listening position [ibid].

2.5. Data Analysis

Following collection and collation of the data for the 23 listeners, it was necessary to assess each listener’s reliability. A data screening procedure was used to ensure that only the data from the most reliable candidates would be included in further data analysis. Listener reliability was based upon the consistency with which the hidden reference was detected. It was assumed that a reliable listener would always grade the hidden reference using the maximum value on the scale. Consequently, any variation from a rating of ‘100’ for the hidden reference was considered an error; the degree of error was determined by the magnitude and frequency of deviation from ‘100’.

The ratings for the hidden reference for each of the three quality attributes were taken into account, but BAQ was given the greatest weight. Whilst it was important to consider the spatial fidelity ratings, the BAQ attribute was considered the best indicator of overall sound quality [5,6].

As a result of the screening process, five listeners were removed due to consistent and/or dramatic errors in their hidden reference ratings; two listeners were removed due to general inconsistency of ratings. This left 16 ‘reliable’ listeners (above the minimum case number of ‘15’) and allowed for robust analysis of variance even when the assumptions of the analyses of variance (ANOVA) are violated (e.g. for non-normal distributions) [9].

2.6. Results

The assumptions of ANOVA analysis were tested and the results indicated it was a reliable analysis method for this experiment.
Three separate ANOVA tests were used to determine the relationships between independent variables (audio excerpt and audio process) and the dependent variables (BAQ, FSF, and SSF ratings). The ANOVA results provided evidence that the audio excerpt variable was not responsible for significant variance of ratings. It was therefore possible to aggregate ratings for the three excerpts. As expected, the audio process was found to be responsible for a significant variance of ratings. The ANOVA results for Experiment I are presented in Appendix A4.

The ‘estimate of effect size’ coefficient, represented by $\eta^2$, is an estimate of the magnitude of significance. This coefficient appears in the ANOVA tables in the column entitled ‘Partial Eta Squared’. The audio excerpt variable was shown to be an insignificant factor for BAQ, FSF and SSF ratings. However, a relationship between the two independent factors (excerpt and process) was shown to be significant for the SSF ratings, because the significance value was approximately 0.03 (estimated size, $\eta^2 = 0.093$). However, this degree of significance was dramatically lower than for the audio processes. Also, the spatial fidelity ratings carried less overall importance. Consequently, this significance was ignored. As a result, the quality ratings were aggregated and analysed together.

The main results of the listening test have been presented in three error bar plots below labelled Figures 4, 5, and 6 for BAQ, FSF, and SSF, respectively. Each plot contains averaged ratings across all three audio excerpts.

The mean BAQ ratings shown in Figure 4 show that the majority of BAQ ratings for Processes A and B were very close to 100. The results of a paired comparison show that differences between mean reference rating and mean rating of Process A and B were insignificant (i.e. the mean difference was less than 5%). It indicates that the 4th and 5th eigenchannels can be completely discarded without causing significant degradation on audio quality.
When 3 out of 5 eigenchannels were removed (Process C), the mean of BAQ ratings was around 85, which was still “good” according to the labels used on the rating scale (see Figure 3). If 4 out of 5 eigenchannels were removed (Process D), large audio quality degradation was caused.

Figures 5 and 6 indicate that effects of eigenchannel removal upon FSF and SSF were generally similar to those of BAQ. However, when 4 out of 5 eigenchannels were removed, the mean BAQ ratings are significantly higher than the FSF and SSF ratings. This suggests that the removal of 4 out of 5 eigenchannels caused more degradation of FSF & SSF than BAQ.

It must be noted that all the foregoing results are valid only for audio material with high interchannel correlations; these results cannot be generalised because only recordings with high correlations were used in this experiment.

2.7. Conclusions

Based on the foregoing discussions, conclusions can be summarised as follows:

1. Degradation of audio quality and spatial fidelity increased with the number of eigenchannels removed.

2. The “less important” eigenchannels (the 4th and 5th eigenchannels) could be completely removed without significantly affecting the audio quality and spatial fidelity for multichannel audio with high interchannel correlation.

3. The KLT eigenchannel removal process resulted in more degradation on spatial fidelity than basic audio quality when most of the eigenchannels were removed.

These findings are very encouraging, showing the potential of KLT for data compression of multichannel audio. However, since only a few audio experts and experimental conditions were used, the first experiment did not provide a thorough understanding of the perceptual importance of KLT eigenchannels. Consequently, a more deliberate experiment was devised in order to investigate the perceptual importance of KLT multichannel audio more thoroughly.

3. EXPERIMENT II

The first aim of the second experiment was to investigate the perceptual effects on audio quality by removing KLT eigenchannels. Furthermore, the nature of the programme material and KLT eigenvalue extraction methods were to be explored as factors which might affect the perceptual effects. Multichannel audio excerpts were selected with a variety of spatial characteristics and different levels of interchannel correlations. In addition, three eigenvalue extraction methods were employed as follows: a covariance matrix based eigenvalue extraction (Cov-KLT), a correlation matrix based eigenvalue extraction (Corr-KLT) and a perceptually optimised eigenvalue extraction method (PO-KLT). Another task of this experiment is to establish the relationship between the order of perceptual importance and the order of statistical importance of KLT eigenchannels. The main research questions of this experiment were as follows:

1. What are the perceptual effects of KLT eigenchannel removal on audio quality as a function of:
   a). the nature of the multichannel audio.
   b). eigenvalue extraction methods.

2. What is the relationship between the order of Perceptual Importance (PI) of eigenchannels and that of the Statistical Importance (SI) of eigenchannels?

3.1. Selection of Audio Excerpts

The audio excerpts for Experiment II were selected from an audio library containing 140 5.1-channel audio excerpts, created by Zielinski and Jiao. The selection was based on the same criteria as Experiment I: interchannel correlation and spatial characteristics.

Whereas only audio materials with high interchannel correlation were selected in Experiment I, in order to investigate how the interchannel correlation influences the performance of KLT process, excerpts with different levels of interchannel correlation were selected for Experiment II. The percentage of
variance represented by the first eigenchannel was used as the measurement of interchannel correlation.

In order to investigate the influence of the spatial characteristics of multichannel audio, selected recordings should cover the four most common spatial scenes: FB, FB+C, FF and FF+C.

Ideally, four groups of excerpts should be chosen according to their spatial scene. Each group should have three recordings with high, medium and low interchannel correlations. But it was found that some of these combinations are very uncommon in the audio library (e.g., low interchannel correlated FB+C recordings). As a result, nine recordings were selected for Experiment II; the details of these recordings are shown in Appendix A5.

3.2. Processing of Audio Material

The processing of the selected audio materials can be seen in Figure 2. The original channels were first transformed using KLT processing; some eigenchannels were then removed before the inverse-KLT was applied. The processes were labelled according to how many eigenchannels remained (see Table 1).

<table>
<thead>
<tr>
<th>Label</th>
<th>Removed eigenchannels</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLT 1</td>
<td>2\textsuperscript{nd}, 3\textsuperscript{rd}, 4\textsuperscript{th}, and 5\textsuperscript{th}</td>
</tr>
<tr>
<td>KLT 2</td>
<td>3\textsuperscript{rd}, 4\textsuperscript{th}, and 5\textsuperscript{th}</td>
</tr>
<tr>
<td>KLT 3</td>
<td>4\textsuperscript{th}, and 5\textsuperscript{th}</td>
</tr>
<tr>
<td>KLT 4</td>
<td>5\textsuperscript{th}</td>
</tr>
</tbody>
</table>

In keeping with the MUSHRA recommendation, the hidden reference was an unprocessed version of the original excerpt and the anchor was obtained by low-pass filtering the original excerpt down to 3.5 kHz in all standard channels. The loudness of all items was equalized.

3.2.1. Eigenvalue extraction methods

In the KLT process, the eigenchannels can be derived based on the eigenvalues and corresponding eigenvectors of either the covariance matrix (Cov-KLT) or the correlation matrix (Corr-KLT) of the originals signals. If the original signals are normalized into a unit scale respectively, for example the range from -1 to +1, and then Cov-KLT processing was applied on the normalised signals, the outputs (eigenchannels) were the same as those created by directly applying Corr-KLT to the original signals. When the units of the original signals are not the same, the Corr-KLT is preferred because it makes the signals comparable [10]. However, in the context of multichannel audio processing in which the signals use the same unit, normalisation will discard the volume differences between channels. The previous study showed that if some channels are much quieter than the others, they could be removed without noticeable effect on the total audio quality [9]. In other words, the perceptual importance of original signals depends greatly upon their volume levels. Therefore, the Corr-KLT was expected to perform less well than the Cov-KLT for some recordings with large volume differences between channels; this needed to be validated through a proper listening test.

Furthermore, it can be argued that two signals with the same volume do not necessarily have the same perceptual importance because human perception is not based on volume but on loudness. The sensitivity of the human auditory system varies across the frequency axis: it is most sensitive around 1-4 kHz and less sensitive at high and low frequencies [11]. Taking this fact into account, a modified eigenvalue extraction method was proposed in this paper, named perceptually optimised eigenvalue extraction method (PO-KLT). The process of PO-KLT is demonstrated in Figure 7. The original signals were filtered by an Inverse Equal Loudness Filter (IELF) then the eigenvalues and eigenvectors were extracted from the covariance matrix of the filtered signals. Based on these eigenvalues and eigenvectors, the KLT process was applied to the originals signals. The eigenchannel removal scheme, shown in Table 1, was then applied before the inverse-KLT.

The design of the IELF was based on the equal loudness contour model shown in Figure 8. The contour of the Minimum Audible Field (MAF) was used. The function of the IELF was used to invert this contour. The amplitude-frequency response of the IELF is shown in Figure 9.
Figure 7: The PO-KLT process

Figure 8: Equal Loudness Contours [12]
3.3. Experiment Design

A MUSHRA listening test was conducted in an ITU-R BS.1116 recommended listening room. The five main loudspeakers were arranged according to the ITU-R BS.775 recommendation. The LFE channel was muted.

Eleven staff and students from the Institute of Sound Recording (IOSR) at the University of Surrey took part in this experiment. All of them were considered experienced listeners.

The listening test included a screening session and three test sessions. The participants were asked to listen to each excerpt and assess the Basic Audio Quality. The graphical user interface is shown in Figure 10. As suggested by [13], a “sort” button was introduced; when desired, the listeners could automatically sort the slider order with this button, such that the differently items were ordered from poorest to the best. The order of the stimuli and pages was randomized.

3.4. Results

As described in Section 3.3, a screening session was carried out before the proper listening test sessions. The inconsistency and discrimination power of the 11 subjects were analysed using the data collected in the screening session. Because one of the eleven listeners had a weak power of discriminating the differences between the reference and the KLT processed items, his or her data was not used in the following analysis.

To answer the first research question of this experiment, the BAQ scores of the four KLT processed items were analysed (the hidden reference and hidden anchor were omitted as they were used solely as control conditions). The results are presented in section 3.4.1. To answer the second research question, the BAQ scores of KLT processed items were first transformed into a form that could be considered an assessment of the perceptual importance of each eigenchannel. The transform methods and the results are presented in Section 3.4.2.
3.4.1. Data analysis: Part I

The assumptions of ANOVA analysis were tested and the results indicated that it was a reliable analysis method for the data.

The ANOVA results of BAQ for the following factors are shown in Appendix A6: eigenvalue extraction method (EXTRAC), recordings (RECORD) and the KLT processing methods (PROCESS). It showed that all the factors and both the 2-way interaction and the 3-way interactions had significant influence on the BAQ. Based on the ANOVA results, the following analyses were carried out and conclusions were drawn.

The mean BAQ for all KLT processing methods and eigenvalue extraction methods for different recordings is presented in Figure 11. Recordings with different levels of interchannel correlation are plotted in different colours and are grouped according to their spatial scenes. It must be mentioned that the data was not balanced for the factor of interchannel correlation.

Figure 10: Graphical user interface for MUSHRA test
Figure 11: Basic Audio Quality for different recordings
(results averaged across all KLT processing methods and eigenvalue extraction methods)

For FB recordings (recording 1, 2 and 3), a clear tendency can be seen that the audio quality degradation caused by the KLT processing methods increased as the degree of interchannel correlation decreased. The same tendency existed in the results for FF and FB+C recordings. Since there was only one FF+C recording involved, the tendency could not be observed. The results indicate that the KLT processing is suitable for recordings with high interchannel correlation better than those with low interchannel correlation.

The audio quality degradations for different eigenvalue extraction methods are compared in Figure 12. It can clearly be seen that, as expected, by increasing the number of removed eigenchannels, the BAQ became worse. Overall, the Cov-KLT and PO-KLT performed similarly and both performed better than the Corr-KLT. However, the result was influenced by KLT processing, as indicated by the ANOVA results. When only one eigenchannel was removed (KLT 4), all of the three eigenvalue extraction methods achieved similar results. Under the conditions KLT 3 and KLT 2 (two and three eigenchannels removed respectively), the Cov-KLT and PO-KLT led to similar degradation but both were better than Corr-KLT. If only one eigenchannel was retained (KLT 1), the PO-KLT became the worst while the Corr-KLT and Cov-KLT had similar performance.

The ANOVA results also indicated that the 3-way interaction of all the factors had a significant influence on BAQ. Therefore, similar graphs were produced for the recordings with different spatial characteristics separately (see Figure 13). It should be noted that the data for the four spatial scenes was not balanced: there were three FB and FF recordings, two FB+C recordings and only one FF+C recordings involved.
Figure 12: Comparison of the performance of three eigenchannels extraction methods on Basic Audio Quality (results averaged across all recordings)

Figure 13: Performance comparison of three eigenchannels extraction methods on Basic Audio Quality as a function of different spatial scene
Since it was found that all three eigenvalue extraction methods achieved similar results under the conditions of KLT1 and KLT 4, only the BAQ resulting from KLT 2 and KLT 3 are considered here. For FB and FB+C recordings, Cov-KLT and PO-KLT led to similar levels of degradation but both were better than Corr-KLT. However, the difference between the performance of PO-KLT and the other two extraction methods increased. For FF and FF+C recordings, however, the three eigenchannel extraction methods performed almost the same.

This was not surprising because Corr-KLT omits the volume difference between channels, as discussed in Section 3.2. For FB and FB+C recordings, the volume difference between front and rear channels is always large. Therefore, Corr-KLT does not efficiently retain the most important information, which therefore leads to bad audio quality. On the other hand, channels of FF and FF+C recordings have relatively comparable volume, which does not make too much difference in the eigenvectors derived from the correlation matrix and covariance matrix. Due to this reason, Corr-KLT performed similarly to the other two methods for FF and FF+C recordings.

However, PO-KLT processing did not achieve significantly better audio quality as was expected.

3.4.2. Data analysis: Part II

In order to answer the second research questions mentioned at the beginning of this section, the measurements of SI and PI of eigenchannels needed to be defined. The SI of certain single eigenchannel was defined as the percentage of total variance which the eigenchannel represents. The values of SI were normalised into a scale from 0 to 100 through multiplication by 100.

As shown in Table 1, KLT 1 only retained the first eigenchannel and KLT 2 retained the 1st and the 2nd eigenchannel. The difference between their BAQ was caused by the absence or presence of the second eigenchannel. Therefore, the difference of BAQ between KLT 1 and KLT 2 could be used as a measurement of the perceptual importance of the 2nd eigenchannel. Generalising this idea, the PI of certain single eigenchannel was calculated from the BAQ scores of the KLT processed items as follows:

\[
\begin{align*}
PI_1 &= BAQ_1; \\
PI_2 &= BAQ_2 - BAQ_1; \\
PI_3 &= BAQ_3 - BAQ_2; \\
PI_4 &= BAQ_4 - BAQ_3; \\
PI_5 &= 100 - BAQ_4; \\
\end{align*}
\]

in which, \( PI_n \) is the perceptual importance of the \( n^{th} \) eigenchannel and the \( BAQ_m \) is the BAQ of the \( m^{th} \) KLT processed item.

The assumptions of ANOVA analysis for these data were tested and the results indicated it was a reliable analysis method.

The ANOVA results of PI for the following factors are shown in Appendix A7: eigenvalue extraction method (EXTRAC), recordings (RECORD) and the eigenchannel number (EC). The ANOVA results showed that the PIs were significantly different for different eigenchannels. The interaction between eigenchannel number and eigenvalue extraction methods, in addition to the interaction between eigenchannel number and the nature of recordings, also had a significant effect on the PIs.

The SI and PI of eigenchannels as for different eigenvalue extraction methods were compared in Figure 14. The results were averaged across all recordings. It can be seen that the SIs were in decreasing orders from the 1st to the 5th eigenchannel for all eigenvalue extraction methods. For the Corr-KLT and Cov-KLT, the PIs had a roughly decreasing order from the 1st to the 5th eigenchannel. Whereas for the PO-KLT, the PIs are in a decreasing order except the PI of the 2nd eigenchannel, which is even higher than the PI of the 1st eigenchannel. In order to check the significant effect, paired comparisons were carried out using a post hoc (Games-Howell) test. The results are shown in Appendix A8.

For the Corr-KLT, \( PI_1 \) is significantly higher than \( PI_2 \) and \( PI_4 \) is significantly higher than \( PI_4 \). The effect between \( PI_2 \) and \( PI_1 \) is not significant; neither is the effect between \( PI_3 \) and \( PI_4 \). For Cov-KLT, similar observations can be seen. It can be concluded that the PI of eigenchannels was in approximately the same order of SI for the Corr-KLT and Cov-KLT. For the PO-KLT, the \( PI_1 \) is significantly higher than other PIs, from which one may conclude that the second eigenchannel seems to be more important than the others.
first eigenchannel. However, it is difficult to generalise these observations as the results strongly depend on programme material, as mentioned above.

The comparison of PI and SI is analysed recording by recording in Appendix A9.

![Comparison between Perceptual Importance (PI) and Statistical Importance (SI) of eigenchannels](image)

**Figure 14:** Comparison between Perceptual Importance (PI) and Statistical Importance (SI) of eigenchannels (results averaged across all recordings)

### 3.5. Discussions

As can be seen from the foregoing analysis, the BAQ of KLT processed items and the perceptual importance of eigenchannels are strongly dependent on programme material. This is because KLT is a signal-dependent transform, in which the transform matrix varies depending on the original signals. The transform matrices of two excerpts are shown in figure 15 as an example.

In the first transform, the first two eigenchannels ($e_1$ and $e_2$) retain little information from the rear channels, whereas in the second matrix, the first two eigenchannels contain much information from the LS channel. It can be expected that these two excerpts will have very different surround spatial fidelity in reproduction using KLT 2 processing (only the 1st and 2nd eigenchannels are retained). However, the difference may not be shown in the results BAQ-based experiment. This encourages some further studies in which more attributes other than BAQ will be assessed.

The proposed perceptual optimised KLT (PO-KLT) methods did not achieve better BAQ than the covariance matrix based KLT (Cov-KLT). A possible explanation is that the BAQ is defined as “all differences between the reference and the evaluated excerpt”, including timbral fidelity, spatial fidelity etc. The PO-KLT weighs the original signals differently than Cov-KLT, because it emphasises the middle frequency content. When eigenchannels are
removed, it may be that the PO-KLT could lead to better timbral fidelity but poor spatial fidelity, which will not necessarily lead to better BAQ overall. However, this also requires further investigation, as timbral fidelity and spatial fidelity was not included in Experiment II.

\[
\begin{bmatrix}
e_1 \\
e_2 \\
e_3 \\
e_4 \\
e_5 \\
\end{bmatrix} =
\begin{bmatrix}
-0.63 & -0.78 & 0.01 & 0.01 & 0.01 & L \\
0.78 & -0.62 & 0.01 & 0.02 & -0.01 & R \\
0.01 & -0.02 & 0.11 & -0.59 & -0.80 & C \\
0.01 & -0.01 & 0.09 & 0.07 & 0.08 & LS \\
0.01 & -0.01 & 0.01 & -0.80 & 0.60 & RS
\end{bmatrix}
\]

Figure 15: KLT matrices for two excerpts

3.6. Conclusions

Conclusions can be summarised as follows:

1. Basic audio quality decreased as the number of removed eigenchannels increased.

2. KLT eigenchannel removal caused less degradation in audio quality for multichannel audio with high interchannel correlation than for those items with low interchannel correlation.

3. Three eigenvalue extraction methods resulted in similar performance of BAQ for FF and FF+C recordings. But for FB and FB+C recordings, the KLT processing based on covariance matrix (Cov-KLT) and perceptually optimised covariance matrix (PO-KLT) caused smaller audio quality degradation than that caused by the KLT processing based on correlation matrix (Corr-KLT).

4. In general, for Corr-KLT and Cov-KLT, the orders of perceptual importance (PI) and statistical importance (SI) of eigenchannels are the same, whereas for the PO-KLT, the perceptual importance of the 2nd eigenchannel is higher than the 1st eigenchannel. However, it is difficult to generalise these observations as the results strongly depend on programme material.

4. SUMMARY

In this paper, the perceptual effects of Karhunen–Löève Transform (KLT) on multichannel audio were systematically studied using two experiments. The first experiment investigated the effects on the audio quality and spatial characteristics caused by removing some KLT eigenchannels. The results showed that some “less important” eigenchannels could be completely removed without significantly affecting audio quality and spatial fidelity for multichannel audio with high interchannel correlation.

The second experiment involved further investigation of the perceptual effects of KLT processing on the audio quality of multichannel audio, as a function of the nature of the multichannel audio and eigenvalue extraction methods. It was found that the KLT eigenchannel removal led to smaller audio quality degradation for high interchannel correlated programme than for those with low interchannel correlation. The KLT processing based on covariance matrix (Cov-KLT) and perceptually optimised covariance matrix (PO-KLT) caused less degradation than processing based on correlation matrix (Corr-KLT) for programme with large interchannel volume differences. The relationship between the order of perceptual importance and the order of statistical importance of KLT eigenchannels was also studied. It was shown that the orders of perceptual importance and statistical importance of eigenchannels are the same using Corr-KLT and Cov-KLT, whereas using the PO-KLT, the perceptual importance of the 2nd eigenchannel is higher than the 1st eigenchannel. However, these observations cannot be simply generalised because they strongly depend on programme material.

Additional attributes beyond basic audio quality will be assessed in future work.

5. ACKNOWLEDGEMENTS

This work was carried out with the financial support of the Engineering and Physical Sciences Research Council, UK (Grant EP/C527100/1).
6. REFERENCES


AES 112th Convention, Munich, Germany, May 10–13, 2002.


APPENDIX

A1. Scree plot of Selected Audio for Experiment I

Figure A2.1: Percentage of variance in each eigenchannel for Excerpt 1.

Figure A2.2: Percentage of variance in each eigenchannel for Excerpt 2.

Figure A2.3: Percentage of variance in each eigenchannel for Excerpt 3.
A2. Selected Audio for Experiment I

<table>
<thead>
<tr>
<th>Excerpt No.</th>
<th>Genre</th>
<th>Spatial Mode</th>
<th>Dialogue</th>
<th>Length</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pop Music</td>
<td>FF</td>
<td>No</td>
<td>20.5 s</td>
<td>Several instruments mixed to all five channels with horns mainly in the rear channels</td>
</tr>
<tr>
<td>2</td>
<td>Movie</td>
<td>FB</td>
<td>Yes</td>
<td>7.7 s</td>
<td>Speech in centre, ambience in all other channels, quiet left and right channels, and very quiet rear channels</td>
</tr>
<tr>
<td>3</td>
<td>Pop Music</td>
<td>FB</td>
<td>No</td>
<td>18.1 s</td>
<td>Instruments mixed mainly to front channels, with reverberations in the rear channels</td>
</tr>
</tbody>
</table>

A3. Mean Differences Between Processes and Reference ($k$-values)

Figure A3.1: Mean differences in dB for Excerpt 1.
Figure A3.2: Mean differences in dB for Excerpt 2.

Figure A3.3: Mean differences in dB for Excerpt 3.
### A4. ANOVA Results for Experiment I

#### Table A4.1: ANOVA results for Basic Audio Quality

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>274237.225</td>
<td>17</td>
<td>16131.602</td>
<td>108.169</td>
<td>.000</td>
<td>.672</td>
</tr>
<tr>
<td>Intercept</td>
<td>1623451.837</td>
<td>1</td>
<td>1623451.837</td>
<td>10885.926</td>
<td>.000</td>
<td>.976</td>
</tr>
<tr>
<td>PROCESS</td>
<td>272047.767</td>
<td>5</td>
<td>54409.553</td>
<td>364.839</td>
<td>.000</td>
<td>.871</td>
</tr>
<tr>
<td>EXCERPT</td>
<td>464.049</td>
<td>2</td>
<td>232.024</td>
<td>1.556</td>
<td>.213</td>
<td>.011</td>
</tr>
<tr>
<td>PROCESS * EXCERPT</td>
<td>1725.410</td>
<td>10</td>
<td>172.541</td>
<td>1.157</td>
<td>.320</td>
<td>.041</td>
</tr>
<tr>
<td>Error</td>
<td>40265.938</td>
<td>270</td>
<td>149.133</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1937955.000</td>
<td>288</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>314503.163</td>
<td>287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .872 (Adjusted R Squared = .864)

#### Table A4.2: ANOVA results for Frontal Spatial Fidelity

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>190954.406</td>
<td>17</td>
<td>11232.612</td>
<td>38.027</td>
<td>.000</td>
<td>.705</td>
</tr>
<tr>
<td>Intercept</td>
<td>1665768.781</td>
<td>1</td>
<td>1665768.781</td>
<td>5639.324</td>
<td>.000</td>
<td>.954</td>
</tr>
<tr>
<td>PROCESS</td>
<td>183312.240</td>
<td>5</td>
<td>36662.448</td>
<td>124.118</td>
<td>.000</td>
<td>.697</td>
</tr>
<tr>
<td>EXCERPT</td>
<td>1543.000</td>
<td>2</td>
<td>771.500</td>
<td>2.612</td>
<td>.075</td>
<td>.019</td>
</tr>
<tr>
<td>PROCESS * EXCERPT</td>
<td>6099.167</td>
<td>10</td>
<td>609.917</td>
<td>2.065</td>
<td>.028</td>
<td>.071</td>
</tr>
<tr>
<td>Error</td>
<td>79751.813</td>
<td>270</td>
<td>295.384</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1936477.000</td>
<td>288</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>270708.219</td>
<td>287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .705 (Adjusted R Squared = .687)

#### Table A4.3: ANOVA results for Surround Spatial Fidelity

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>311747.069</td>
<td>17</td>
<td>18338.063</td>
<td>63.869</td>
<td>.000</td>
<td>.801</td>
</tr>
<tr>
<td>Intercept</td>
<td>1384170.681</td>
<td>1</td>
<td>1384170.681</td>
<td>4820.888</td>
<td>.000</td>
<td>.947</td>
</tr>
<tr>
<td>PROCESS</td>
<td>301952.569</td>
<td>5</td>
<td>60390.514</td>
<td>210.332</td>
<td>.000</td>
<td>.796</td>
</tr>
<tr>
<td>EXCERPT</td>
<td>1614.361</td>
<td>2</td>
<td>907.181</td>
<td>3.160</td>
<td>.044</td>
<td>.023</td>
</tr>
<tr>
<td>PROCESS * EXCERPT</td>
<td>7980.139</td>
<td>10</td>
<td>798.014</td>
<td>2.779</td>
<td>.003</td>
<td>.093</td>
</tr>
<tr>
<td>Error</td>
<td>77522.250</td>
<td>270</td>
<td>287.119</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>173440.000</td>
<td>288</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>389269.319</td>
<td>287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .801 (Adjusted R Squared = .788)
## A5. Selected Audio Experts for Experiment II

<table>
<thead>
<tr>
<th>Excerpt No.</th>
<th>Spatial Scene</th>
<th>Level of Interchannel Correlation (the percentage of variance represented by the first eigenchannel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FF</td>
<td>High (93%)</td>
</tr>
<tr>
<td>2</td>
<td>FB</td>
<td>Medium (58%)</td>
</tr>
<tr>
<td>3</td>
<td>FB</td>
<td>Low (33%)</td>
</tr>
<tr>
<td>4</td>
<td>FF</td>
<td>High (75%)</td>
</tr>
<tr>
<td>5</td>
<td>FF</td>
<td>Medium (52%)</td>
</tr>
<tr>
<td>6</td>
<td>FF</td>
<td>Low (41%)</td>
</tr>
<tr>
<td>7</td>
<td>FB+C</td>
<td>High (93%)</td>
</tr>
<tr>
<td>8</td>
<td>FB+C</td>
<td>Medium (58%)</td>
</tr>
<tr>
<td>9</td>
<td>FF+C</td>
<td>High (33%)</td>
</tr>
</tbody>
</table>
A6. ANOVA results for the basic audio quality (BAQ) of KLT processed items for all listeners

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>60378.9485²</td>
<td>107</td>
<td>5642.892</td>
<td>26.734</td>
<td>.000</td>
<td>.746</td>
</tr>
<tr>
<td>Intercept</td>
<td>5000722.31</td>
<td>1</td>
<td>5000722.31</td>
<td>23691.303</td>
<td>.000</td>
<td>.961</td>
</tr>
<tr>
<td>EXTRAC</td>
<td>5921.124</td>
<td>2</td>
<td>2960.562</td>
<td>14.026</td>
<td>.000</td>
<td>.028</td>
</tr>
<tr>
<td>RECORD</td>
<td>30833.869</td>
<td>8</td>
<td>3854.234</td>
<td>18.260</td>
<td>.000</td>
<td>.131</td>
</tr>
<tr>
<td>PROCESS</td>
<td>480272.04</td>
<td>3</td>
<td>160090.668</td>
<td>758.442</td>
<td>.000</td>
<td>.701</td>
</tr>
<tr>
<td>EXTRAC * RECORD</td>
<td>24297.643</td>
<td>16</td>
<td>1518.603</td>
<td>7.194</td>
<td>.000</td>
<td>.106</td>
</tr>
<tr>
<td>EXTRAC * PROCESS</td>
<td>12934.380</td>
<td>6</td>
<td>2155.730</td>
<td>10.213</td>
<td>.000</td>
<td>.059</td>
</tr>
<tr>
<td>RECORD * PROCESS</td>
<td>21740.213</td>
<td>24</td>
<td>905.842</td>
<td>4.291</td>
<td>.000</td>
<td>.096</td>
</tr>
<tr>
<td>EXTRAC * RECORD * PROCESS</td>
<td>27790.254</td>
<td>48</td>
<td>578.964</td>
<td>2.743</td>
<td>.000</td>
<td>.119</td>
</tr>
<tr>
<td>Error</td>
<td>205168.200</td>
<td>972</td>
<td>211.078</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>580968.00</td>
<td>1080</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>808957.685</td>
<td>1079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .746 (Adjusted R Squared = .718)

A7. ANOVA Results for the perceptual importance (PI) of KLT eigenchannels for all listeners

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>225248.692²</td>
<td>107</td>
<td>2105.128</td>
<td>11.677</td>
<td>.000</td>
<td>.562</td>
</tr>
<tr>
<td>Intercept</td>
<td>287630.208</td>
<td>1</td>
<td>287630.208</td>
<td>1595.446</td>
<td>.000</td>
<td>.621</td>
</tr>
<tr>
<td>EXTRAC</td>
<td>890.906</td>
<td>2</td>
<td>445.453</td>
<td>2.471</td>
<td>.085</td>
<td>.005</td>
</tr>
<tr>
<td>RECORD</td>
<td>3595.650</td>
<td>8</td>
<td>449.456</td>
<td>2.493</td>
<td>.011</td>
<td>.020</td>
</tr>
<tr>
<td>EC</td>
<td>90357.418</td>
<td>3</td>
<td>30119.139</td>
<td>167.067</td>
<td>.000</td>
<td>.340</td>
</tr>
<tr>
<td>EXTRAC * RECORD</td>
<td>1516.761</td>
<td>16</td>
<td>94.798</td>
<td>.526</td>
<td>.935</td>
<td>.009</td>
</tr>
<tr>
<td>EXTRAC * EC</td>
<td>22689.124</td>
<td>6</td>
<td>3781.521</td>
<td>20.976</td>
<td>.000</td>
<td>.115</td>
</tr>
<tr>
<td>RECORD * EC</td>
<td>43782.024</td>
<td>24</td>
<td>1824.251</td>
<td>10.119</td>
<td>.000</td>
<td>.200</td>
</tr>
<tr>
<td>EXTRAC * RECORD * EC</td>
<td>62416.809</td>
<td>48</td>
<td>1300.350</td>
<td>7.213</td>
<td>.000</td>
<td>.263</td>
</tr>
<tr>
<td>Error</td>
<td>175234.100</td>
<td>972</td>
<td>180.282</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>688113.000</td>
<td>1080</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>400482.792</td>
<td>1079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .562 (Adjusted R Squared = .514)
A8. Post hoc test of perceptual importance

Table. A8.1: Post hoc test of perceptual importance for the Corr-KLT

Multiple Comparisons\textsuperscript{a}

<table>
<thead>
<tr>
<th>(i) eigen channel number</th>
<th>(j) eigen channel number</th>
<th>Mean Difference (i–j)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC1</td>
<td>EC2</td>
<td>16.7333*</td>
<td>2.7452</td>
<td>.000</td>
<td>9.1611 – 24.3055</td>
</tr>
<tr>
<td>EC1</td>
<td>EC3</td>
<td>24.1222*</td>
<td>2.86773</td>
<td>.000</td>
<td>16.2188 – 32.0256</td>
</tr>
<tr>
<td>EC1</td>
<td>EC4</td>
<td>20.0556*</td>
<td>3.24690</td>
<td>.000</td>
<td>11.1016 – 29.0095</td>
</tr>
<tr>
<td>EC1</td>
<td>EC5</td>
<td>28.1444*</td>
<td>2.29214</td>
<td>.000</td>
<td>21.8106 – 34.4783</td>
</tr>
<tr>
<td>EC2</td>
<td>EC1</td>
<td>-16.7333*</td>
<td>2.7452</td>
<td>.000</td>
<td>-24.3055 – -9.1611</td>
</tr>
<tr>
<td>EC2</td>
<td>EC3</td>
<td>7.3889</td>
<td>2.80128</td>
<td>.068</td>
<td>-3.319 – 15.1096</td>
</tr>
<tr>
<td>EC2</td>
<td>EC4</td>
<td>3.3222</td>
<td>3.18837</td>
<td>.835</td>
<td>-5.4726 – 12.1170</td>
</tr>
<tr>
<td>EC2</td>
<td>EC5</td>
<td>11.4111*</td>
<td>2.20844</td>
<td>.000</td>
<td>5.3109 – 17.5113</td>
</tr>
<tr>
<td>EC3</td>
<td>EC1</td>
<td>-24.1222*</td>
<td>2.86773</td>
<td>.000</td>
<td>-32.0256 – -16.2188</td>
</tr>
<tr>
<td>EC3</td>
<td>EC2</td>
<td>-7.3889</td>
<td>2.80128</td>
<td>.068</td>
<td>-15.1096 – 3.319</td>
</tr>
<tr>
<td>EC3</td>
<td>EC4</td>
<td>-4.0667</td>
<td>3.29252</td>
<td>.731</td>
<td>-13.1449 – 5.0116</td>
</tr>
<tr>
<td>EC3</td>
<td>EC5</td>
<td>4.0222</td>
<td>2.35632</td>
<td>.433</td>
<td>-2.4908 – 10.5352</td>
</tr>
<tr>
<td>EC4</td>
<td>EC1</td>
<td>-20.0556*</td>
<td>3.24690</td>
<td>.000</td>
<td>-29.0095 – -11.1016</td>
</tr>
<tr>
<td>EC4</td>
<td>EC3</td>
<td>4.0667</td>
<td>3.29252</td>
<td>.731</td>
<td>-5.0116 – 13.1449</td>
</tr>
<tr>
<td>EC4</td>
<td>EC5</td>
<td>8.0889*</td>
<td>2.80548</td>
<td>.037</td>
<td>.3211 – 15.8567</td>
</tr>
<tr>
<td>EC5</td>
<td>EC1</td>
<td>-28.1444*</td>
<td>2.29214</td>
<td>.000</td>
<td>-34.4783 – -21.8106</td>
</tr>
<tr>
<td>EC5</td>
<td>EC2</td>
<td>-11.4111*</td>
<td>2.20844</td>
<td>.000</td>
<td>-17.5113 – -5.3109</td>
</tr>
<tr>
<td>EC5</td>
<td>EC3</td>
<td>-4.0222</td>
<td>2.35632</td>
<td>.433</td>
<td>-10.5352 – 2.4908</td>
</tr>
<tr>
<td>EC5</td>
<td>EC4</td>
<td>-8.0889*</td>
<td>2.80548</td>
<td>.037</td>
<td>-15.8567 – -3.211</td>
</tr>
</tbody>
</table>

Based on observed means.
\textsuperscript{a}. The mean difference is significant at the .05 level.
\textsuperscript{a}. Eigenchannel extraction methods = Corr–KLT
Table A8.2: Post hoc test of perceptual importance for the Cov-KLT

Multiple Comparisons

Games–Howell

<table>
<thead>
<tr>
<th>(i) eigen channel number</th>
<th>(j) eigen channel number</th>
<th>Mean Difference (i–j)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>EC 1</td>
<td>EC 2</td>
<td>2.3667</td>
<td>2.82508</td>
<td>.919</td>
<td>-5.4191</td>
</tr>
<tr>
<td>EC 1</td>
<td>EC 3</td>
<td>24.2778*</td>
<td>2.48742</td>
<td>.000</td>
<td>17.4155</td>
</tr>
<tr>
<td>EC 1</td>
<td>EC 4</td>
<td>29.6000*</td>
<td>2.38435</td>
<td>.000</td>
<td>23.0161</td>
</tr>
<tr>
<td>EC 1</td>
<td>EC 5</td>
<td>27.4222*</td>
<td>2.31493</td>
<td>.000</td>
<td>21.0245</td>
</tr>
<tr>
<td>EC 2</td>
<td>EC 1</td>
<td>-2.3667</td>
<td>2.82508</td>
<td>.919</td>
<td>-10.1525</td>
</tr>
<tr>
<td>EC 2</td>
<td>EC 3</td>
<td>21.9111*</td>
<td>2.45298</td>
<td>.000</td>
<td>15.1446</td>
</tr>
<tr>
<td>EC 2</td>
<td>EC 4</td>
<td>27.2333*</td>
<td>2.34841</td>
<td>.000</td>
<td>20.7497</td>
</tr>
<tr>
<td>EC 2</td>
<td>EC 5</td>
<td>25.0556*</td>
<td>2.27789</td>
<td>.000</td>
<td>18.7613</td>
</tr>
<tr>
<td>EC 3</td>
<td>EC 1</td>
<td>-24.2778*</td>
<td>2.48742</td>
<td>.000</td>
<td>-31.1401</td>
</tr>
<tr>
<td>EC 3</td>
<td>EC 4</td>
<td>5.3222*</td>
<td>1.92904</td>
<td>.050</td>
<td>.0049</td>
</tr>
<tr>
<td>EC 3</td>
<td>EC 5</td>
<td>3.1444</td>
<td>1.84254</td>
<td>.433</td>
<td>-1.9366</td>
</tr>
<tr>
<td>EC 4</td>
<td>EC 1</td>
<td>-29.6000*</td>
<td>2.38435</td>
<td>.000</td>
<td>-36.1839</td>
</tr>
<tr>
<td>EC 4</td>
<td>EC 2</td>
<td>-27.2333*</td>
<td>2.34841</td>
<td>.000</td>
<td>-33.7170</td>
</tr>
<tr>
<td>EC 4</td>
<td>EC 3</td>
<td>-5.3222*</td>
<td>1.92904</td>
<td>.050</td>
<td>-10.6395</td>
</tr>
<tr>
<td>EC 4</td>
<td>EC 5</td>
<td>-2.1778</td>
<td>1.70084</td>
<td>.704</td>
<td>-6.8658</td>
</tr>
<tr>
<td>EC 5</td>
<td>EC 1</td>
<td>-27.4222*</td>
<td>2.31493</td>
<td>.000</td>
<td>-33.8199</td>
</tr>
<tr>
<td>EC 5</td>
<td>EC 2</td>
<td>-25.0556*</td>
<td>2.27789</td>
<td>.000</td>
<td>-31.3498</td>
</tr>
<tr>
<td>EC 5</td>
<td>EC 3</td>
<td>-3.1444</td>
<td>1.84254</td>
<td>.433</td>
<td>-8.2255</td>
</tr>
<tr>
<td>EC 5</td>
<td>EC 4</td>
<td>2.1778</td>
<td>1.70084</td>
<td>.704</td>
<td>-2.5102</td>
</tr>
</tbody>
</table>

Based on observed means.

* The mean difference is significant at the .05 level.

a. Eigenchannel extraction methods = Cov-KLT
Table A8.3: Post hoc test of perceptual importance for the PO-KLT

**Multiple Comparisons**

Dependent Variable: perceptual importance

Games–Howell

<table>
<thead>
<tr>
<th>(I) eigen channel number</th>
<th>(J) eigen channel number</th>
<th>Mean Difference (I–J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval Lower Bound</th>
<th>95% Confidence Interval Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC1</td>
<td>EC2</td>
<td>-10.8111*</td>
<td>3.21574</td>
<td>.008</td>
<td>-19.6739</td>
<td>-1.9483</td>
</tr>
<tr>
<td></td>
<td>EC3</td>
<td>16.7111*</td>
<td>2.6373</td>
<td>.000</td>
<td>9.3600</td>
<td>24.0622</td>
</tr>
<tr>
<td></td>
<td>EC4</td>
<td>20.3667*</td>
<td>2.68492</td>
<td>.000</td>
<td>12.9582</td>
<td>27.7751</td>
</tr>
<tr>
<td></td>
<td>EC5</td>
<td>21.8444*</td>
<td>2.42090</td>
<td>.000</td>
<td>15.1433</td>
<td>28.5455</td>
</tr>
<tr>
<td>EC2</td>
<td>EC1</td>
<td>10.8111*</td>
<td>3.21574</td>
<td>.008</td>
<td>1.9483</td>
<td>19.6739</td>
</tr>
<tr>
<td></td>
<td>EC3</td>
<td>27.5222*</td>
<td>2.78500</td>
<td>.000</td>
<td>19.8331</td>
<td>35.2114</td>
</tr>
<tr>
<td></td>
<td>EC4</td>
<td>31.1778*</td>
<td>2.80527</td>
<td>.000</td>
<td>23.4340</td>
<td>38.9216</td>
</tr>
<tr>
<td></td>
<td>EC5</td>
<td>32.6556*</td>
<td>2.55373</td>
<td>.000</td>
<td>25.5833</td>
<td>39.7278</td>
</tr>
<tr>
<td>EC3</td>
<td>EC1</td>
<td>-16.7111*</td>
<td>2.6373</td>
<td>.000</td>
<td>-24.0622</td>
<td>-9.3600</td>
</tr>
<tr>
<td></td>
<td>EC2</td>
<td>-27.5222*</td>
<td>2.78500</td>
<td>.000</td>
<td>-35.2114</td>
<td>-19.8331</td>
</tr>
<tr>
<td></td>
<td>EC4</td>
<td>3.6556</td>
<td>2.15035</td>
<td>.437</td>
<td>-2.2707</td>
<td>9.5818</td>
</tr>
<tr>
<td></td>
<td>EC5</td>
<td>5.1333*</td>
<td>1.80998</td>
<td>.041</td>
<td>1.1378</td>
<td>10.1288</td>
</tr>
<tr>
<td>EC4</td>
<td>EC1</td>
<td>-20.3667*</td>
<td>2.68492</td>
<td>.000</td>
<td>-27.7751</td>
<td>-12.9582</td>
</tr>
<tr>
<td></td>
<td>EC2</td>
<td>-31.1778*</td>
<td>2.80527</td>
<td>.000</td>
<td>-38.9216</td>
<td>-23.4340</td>
</tr>
<tr>
<td></td>
<td>EC3</td>
<td>-3.6556</td>
<td>2.15035</td>
<td>.437</td>
<td>-9.5818</td>
<td>2.2707</td>
</tr>
<tr>
<td></td>
<td>EC5</td>
<td>1.4778</td>
<td>1.84102</td>
<td>.929</td>
<td>-3.6042</td>
<td>6.5598</td>
</tr>
<tr>
<td>EC5</td>
<td>EC1</td>
<td>-21.8444*</td>
<td>2.42090</td>
<td>.000</td>
<td>-28.5455</td>
<td>-15.1433</td>
</tr>
<tr>
<td></td>
<td>EC2</td>
<td>-32.6556*</td>
<td>2.55373</td>
<td>.000</td>
<td>-39.7278</td>
<td>-25.5833</td>
</tr>
<tr>
<td></td>
<td>EC3</td>
<td>-5.1333*</td>
<td>1.80998</td>
<td>.041</td>
<td>-10.1288</td>
<td>-1.378</td>
</tr>
<tr>
<td></td>
<td>EC4</td>
<td>-1.4778</td>
<td>1.84102</td>
<td>.929</td>
<td>-6.5598</td>
<td>3.6042</td>
</tr>
</tbody>
</table>

Based on observed means.

* The mean difference is significant at the .05 level.

a. Eigenchannel extraction methods = PO–KLT
A9. Comparison between PI and SI of eigenchannels presented separately for each 9 recordings

Figure A9.1: Comparison between PI and SI of eigenchannels for the Corr-KLT presented separately for each 9 recordings
Figure A9.2: Comparison between PI and SI of eigenchannels for the Cov-KLT presented separately for each 9 recordings.
Figure A9.3: Comparison between PI and SI of eigenchannels for the PO-KLT presented separately for each 9 recordings