Analysis and automatic annotation of singer's postures during concert and rehearsal

Maarten Grachten **Michiel Demey** **Dirk Moelants** **Marc Leman**
Institute of Psychoacoustics and Electronic Music (IPEM)
Department of Musicology – Ghent University
Ghent, Belgium
http://www.ipem.ugent.be

ABSTRACT

Bodily movement of music performers is widely acknowledged to be a means of communication with the audience. For singers, where the necessity of movement for sound production is limited, postures, i.e. static positions of the body, may be relevant in addition to actual movements. In this study, we present the results of an analysis of a singer's postures, focusing on differences in postures between a dress rehearsal without audience and a concert with audience. We provide an analysis based on manual annotation of postures and propose and evaluate methods for automatic annotation of postures based on motion sensing data, showing that automatic annotation is a viable alternative to manual annotation. Results furthermore suggest that the presence of an audience leads the singer to use more 'open' postures, and differentiate more between different postures. Also, speed differences of transitions from one posture to another are more pronounced in concert than during rehearsal.

1. INTRODUCTION AND RELATED WORK

The performance of music is naturally accompanied by corporal gestures of the performing musician. The form and roles of these gestures in music performance appear to be heterogeneous and have led to considerable investigation [1, 2, 3]. A distinction has been made between four categories of gestures: (a) sound-producing gestures, necessary to create sound; (b) communicative gestures between musicians or between the musician and the audience; (c) sound-facilitating gestures, which accompany the first category; and (d) sound accompanying gestures like dancing, that are generally not produced by the musician himself [4, 5]. In singing performance sound-producing gestures are very limited, the actual sound is produced inside the body and the only visually perceivable elements are the articulation of the mouth and possibly the breathing. Sound-facilitating gestures are equally limited and restricted to posture changes that facilitate singing in e.g. high or low registers. Most of the gestures perceived in singing performance can therefore be attributed to communication, and more specifically to communication with the audience. It can be shown through subtle gestures like facial expressions, but is most obvious in the movements of hand and arms and to a lesser extend the head and the upper body. These gestures may for example reflect the temporal structure of the piece, convey the mood of the piece, or underline the meaning or importance of certain words or phrases. Therefore the study of gestural aspects of singing performance can give us information on structural elements of the music, as well as on expression and communication.

Many studies on perceived expressivity in musician's gestures exist, based on direct measurements as well as on observations (for an overview see [5]), however these studies almost entirely deal with instrumentalists. Despite the importance of gestural communication in singing performance, studies on this topic are scarce. If we want to study gestural communication between a singer and the public, we can compare a performance in a dress rehearsal and in the actual concert. The performance during the dress rehearsal is supposed to be technically and interpretatively mature and will be performed in the same setting and the same order as the concert. The one aspect that is clearly different is the presence of a public. In general performing musicians acknowledge that the interaction with the public affects their performance, but very little is known about what is actually changing and how. Intuitively one could say that musical elements like the timing and dynamics change, but also the gestural communication is changing, using different movements, facial expressions or eye contact. By using multi-modal measurements (audio, video, movement sensors) and the development of new analytical techniques, we can quantify different aspects of the performance and thus develop a set of parameters that can be used to compare performances, in casu to detect the differences between a rehearsal and a concert performance.

This paper focuses on a singer's postures and the transitions between them. Techniques that allow automatic classification of a set of typical postures are presented and applied to the comparison of the recordings from the dress rehearsal and the concert.

2. DATA

A dress rehearsal and a concert by singer Chia-Fen Wu and viola da gamba player Dirk Moelants were recorded. The

Index	Composer	Piece
01	Giulio Caccini	Dolcissimo Sospiri
02	Giulio Caccini	Movetevi a pieta
03	Barbara Strozzi	Moralit amorosa
04	Barbara Strozzi	Non occore
05	Richard Sumarte	Daphne
06	John Dowland	Come Again
07	John Dowland	Flow my tears
08	Robert Johnson	Hark, hark, the lark
09	Tobias Hume	Tobacco
10	Thomas Morley	It was a lover and his lass
11	Richard Sumarte	What if a day
12	Richard Sumarte	Whoope doe me no harme
13	Henry Purcell	How sweet it is to love
14	Henry Purcell	Music for a while
15	Henry Purcell	If music be the food of love
16	Teng Yu-Hsien	Bang Chun Hong
17	Yang San-Lang	Go Luan Hue
18	traditional Chinese	Ye Lai Shiang

Table 1. Overview of the concert program analyzed in this paper. The pieces will henceforth be referred to by the numbers on the left

program is given in table 1. Three pieces (05, 11 & 12) are short pieces for solo viola da gamba, they will not be considered in the present study. All the other pieces are performed by a (soprano) voice with viola da gamba accompaniment. The first 15 pieces are period-style arrangements of 17th century baroque music. The last piece (18) is a traditional Chinese song, in the concert performance it was brought as an encore. The two pieces before (16 & 17) are Taiwanese art songs from the middle of the previous century.

Three different measurements were made of the two performances: an audio recording, a measurement of the movement and a video recording. The audio was recorded using a mobile recorder with a built-in microphone (Zoom H2) positioned at the side of the stage. The movement of the performers was measured using wireless accelerometers with a range of +/-3g and with 2 or 3 sensitive axes. The singer had a sensor on each wrist and one sensor on her back. The sensors were attached to the skin with medical bandage tape underneath the clothes in such a way that they did not hamper the movements of the performers and that they were not visible for the audience. The accelerometers were connected to a standalone, battery powered, wireless ADC module (Wi-microDig, Infusion Systems) that digitizes the analogue sensor data and transmits this data wireless via Bluetooth. A Bluetooth class 1 interface was used enabling a range of 100m making it possible to collect the data from the balcony in the back of the concert hall. The sensor data was recorded at a sampling rate of 100Hz using a Max/MSP patch. Furthermore, the entire concert and rehearsal was videotaped using a Canon HV30 camera.

3. DATA ANALYSIS

3.1 Posture occurrences and durations

A first step in the analysis was the creation of a ground-truth of the postures used by the singer. First all the video material was watched to determine the different postures

Figure 1. Typical examples of the 10 different categories of postures as found in the singers performance

used. In total 10 categories of postures could be distinguished: 0. Arms down, hands joined, 1. hands joined in front of the body, 2. both arms slightly spread in front of the body, 3. both arms in front of the body, left arm above right arm, 4. both arms in front of the body, right arm above left arm, 5. right arm next to the body, left arm in front, 6. left arm next to the body, right arm in front, 7. two arms next to the body, 8. one arm next to the body, the other spread open and 9. two arms spread out. Examples of each posture, taken from the video recording are shown in figure 1.

In a next stage these 10 postures were used in a detailed manual analysis done by author DM using the program Annotation (`http://www.saysosoft.com`). A global overview of the postures used in the rehearsal and the concert is given in table 2. This shows that the singer changed posture more often during the concert, reflected in a increase of total postures of 9.75%. Postures 0, 8 and 9 are not very common, but still it is striking that they only occur during the concert performance. In figure 2 the representation is a bit simplified by grouping these three gestures as ‘others’ and by adding the symmetrical postures 3/4 and 5/6 together. The upper two pie diagrams represent the number of postures counted (analogous to table 2), the lower two represent the total time of each posture type. It shows that the most important change occurs in the number of and time spent in posture 1, that is hands joined in front of the body. This posture is considered as a ‘resting’ or ‘starting’ posture, and can thus be seen as the least ‘communicative’ or ‘expressive’ posture. This posture is largely replaced by more ‘open’ postures in the concert performance (types 5, 6 and 7).

As we count more stable postures in the concert performance, this implies that there are more transitions. In total the number of transitions increases with 10.26%, from 341 to 376. Despite the larger number of transitions, we see that the average duration of a transition increases with

posture	rehearsal	concert
0	0	3
1	57	27
2	65	71
3	37	31
4	3	1
5	36	86
6	43	26
7	118	141
8	0	1
9	0	7
sum	359	394

Table 2. Comparison of the number of posture occurrences for each of the 10 categories in rehearsal and concert performance, summed over all pieces

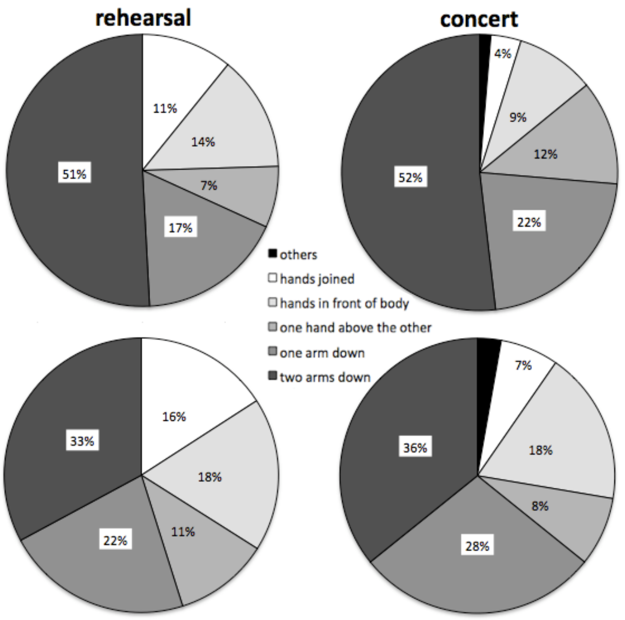

Figure 2. Top row: The relative occurrence of each posture (number of instances); Bottom row: Cumulative duration of each posture

6.45%, from 1.61s to 1.72s. However, as the distribution of the posture types in both performances is different (cf. supra), it is dangerous to make such a general comparison. Therefore the 20 most important transitions, occurring at least twice in each performance where selected. The average length of these transitions in both performances is shown in figure 3. It shows that there is often a large increase in average duration, while only four transition types show a small decrease in average duration. The grand average of these 20 transitions increases with 21.62% from 1.54s to 1.87s between rehearsal and concert performance. This shows that the singer puts more emphasis on the transitions, by making them slower or by increasing the distance.

3.2 Posture and transition analysis

For a further quantification of the postures by the singer the 3D accelerometer sensor data is analyzed. Starting from the 3D acceleration measured at the wrists of the singer it

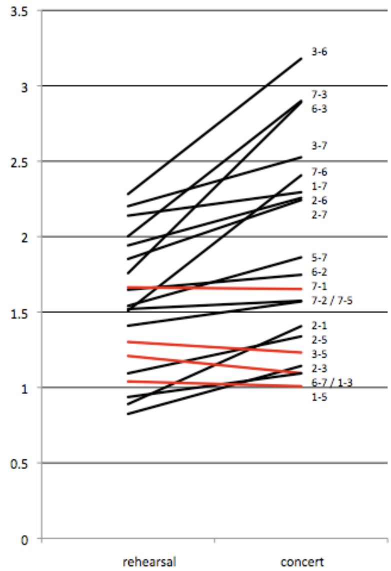

Figure 3. Differences in transition durations between rehearsal and concert, for the 20 most frequent transitions between postures

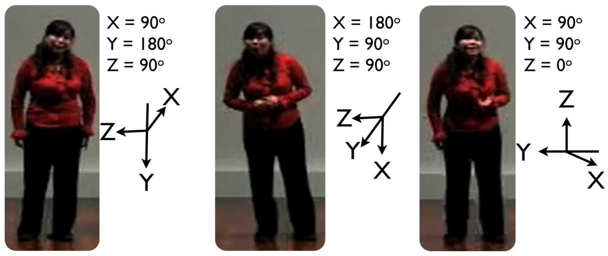

Figure 4. Three postures with the orientation of the 3D accelerometer on the left arm

is possible to calculate the orientation of the forearm with respect to the vertical direction (along the gravity force). This calculation is valid for slow movements in which the total size of the acceleration stays around the value of 1 g. The angles are determined with the following formula:

$$\alpha_i = \arccos\left[\frac{a_i}{\sqrt{a_x^2 + a_y^2 + a_z^2}}\right] \text{ with } i = x, y, z$$

where a_x, a_y, a_z are the 3 components of the acceleration and α_i is the angle between the vertical direction and the acceleration direction under study. In figure 4 these angles are illustrated for the left arm of the singer for three typical postures.

For the study described below the angle in the direction along the forearm is studied (the y-component of the accelerometer). Given the manual annotations of the different postures of the singer the angles for left and right forearm (denoted α_l and α_r respectively) can be determined from the accelerometer data. This leads to the plots shown in figure 5.

One can see that there is a large overlap between posture 1 and 2 which both have similar angles and cannot be disentangled from each other. Furthermore the angular spread

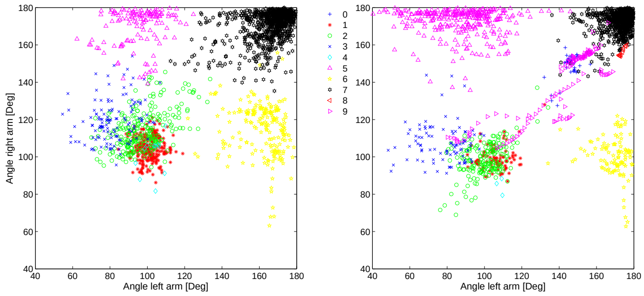

Figure 5. The angle of the left and right forearm extracted from accelerometer data for the rehearsal (left) and concert (right) condition. The different colors/symbols represent the different postures determined from manual annotation

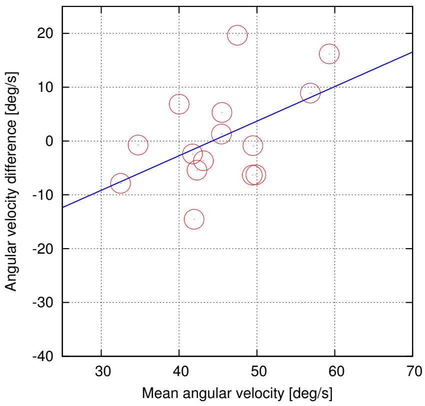

Figure 6. The correlation between the mean angular velocity and the difference in angular velocity for concert and rehearsal conditions. Each data point corresponds to the transitions occurring in a single song of the repertoire

of the data is smaller in the rehearsal condition. This can be quantified by determining the angular center of each posture. The mean Euclidean distance between these centers (in degrees) for the rehearsal and concert conditions are $42.91°$, and $50.72°$, respectively, with a mean difference of $7.81°$. A (single-sided) Wilcoxon signed rank test between the angle pairs corresponding to each posture reveals that the distances between postures are significantly smaller during rehearsal than during the concert condition ($z = -2.5732, p = 0.005$).

On the other hand we can also study the transitions between postures based on the calculated angles. When looking at the angular velocity of the transitions occurring in each song a clear correlation is found between the mean velocity of the two conditions and the difference between the velocities as can be seen in figure 6. A linear fit shows a slope of .64. Note that the data point with the highest difference corresponds to the last song which was brought as an encore.

These results are in accordance with previous findings stated in [6] where an increase in intensity of movement was found in the concert condition for songs with a higher average value.

3.3 Automatic recognition of postures during rehearsal and concert

The above analysis of the data is based on manual annotations of postures using video recordings of the performances. This is generally considered to be the most reliable method of annotating data, but even for a moderate amount of data such as used in this study, manual annotation is a very laborious task. However, it is to be expected that each of the postures identified above should have its own signature in the forearm angles as computed from the acceleration data. Indeed, plotting the left and right forearm angles corresponding to the different postures (figure 5) shows various clearly identifiable clusters.

In this subsection we describe a straight-forward method to cluster pairs of forearm angles into postures, and discuss the results. The method consists in first separating postures and transitions, and subsequently clustering the data corresponding postures. In the third part of this subsection we describe the results of the automatic annotation as applied to the data, and evaluate them using the manual annotation.

3.3.1 Separation of postures and transitions

The first step is to determine which time segments correspond to postures and which to transitions from one posture into another. Since a posture is by definition a more or less static position of the body, we use a criterion that states that a time segment represents a posture precisely if the average change of angle per arm does not exceed a particular threshold γ within that segment. This corresponds to the following criterion:

$$\max \begin{bmatrix} \frac{1}{(K-1)} \sum_{n=2}^{K} |\alpha_l(n) - \alpha_l(n-1)|, \\ \frac{1}{(K-1)} \sum_{n=2}^{K} |\alpha_r(n) - \alpha_r(n-1)| \end{bmatrix} < \gamma \quad (1)$$

where K is the window size, $\alpha_l(n)$, and $\alpha_r(n)$ are the angles of left and right forearm at time n respectively, and γ is the threshold parameter, representing the maximum average change of angle allowed in a posture (in degrees per second).

3.3.2 Clustering of postures

Literature on machine learning and data mining offers a myriad of clustering algorithms. Most, if not all of these algorithms in some form or another rely on parameters to reflect information about the data to be clustered. Depending on the algorithm, parameters can reflect for instance the number of clusters to identify, the maximal variance tolerated within a cluster, or the a priori probability that two data points belong to the same cluster.

In our case, we can easily obtain useful knowledge about the data to be clustered, in the form of approximate forearm angles for the singer's postures. A set of postures is identified by skimming over the video recording, and for each posture an estimation of the forearm angles is made. For example, for the posture 'right hand above left hand' (in front of the body), we expect forearm angles approximately in the middle of the range $180°$ (arm straight down) and $0°$ (arm straight up), where the angle of the right arm is somewhat smaller than that of the left arm. In this manner,

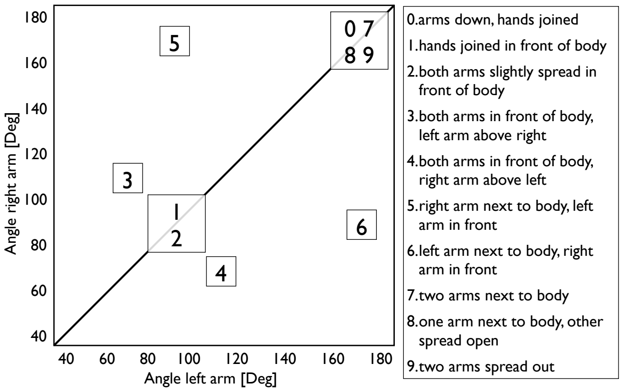

Figure 7. Postures and their typical left and right forearm angle configurations; the x and y axis display the angles of the left and right forearm in degrees, respectively

pairs of forearm angles are defined to serve as cluster prototypes. The prototypes for each posture are displayed in figure 7. Note that the figure contains six postures, rather than ten. The reason for this simplification is that in the limited representation of postures as angles of left and right forearms, some postures are not distinguishable. In particular, postures 1 and 2 (hands joined, and hands slightly open in front of the body), and the various postures where both arms are stretched downward (postures 8, 9, 7, and 0), are all characterized by very similar angles. For this reason, posture 2 has been merged into posture 1 (hands joined), and postures 8, 9, and 0 have been merged into the more frequent posture 7 (both hands down).

Because of the availability of cluster prototypes as a form of knowledge about the data, our particular interest is in clustering algorithms that can take advantage of this knowledge. One such algorithm is the well-known K-means algorithm. Typically, the algorithm is initialized by choosing K random cluster prototypes, but by setting the cluster prototypes deliberately, we can make the algorithm start from hypothesized prototypes, and update these prototypes in accordance with the data.

We also present another simple clustering algorithm that we designed for the purpose of adapting hypothesized cluster prototypes. The algorithm starts from a given set of prototypes, and assigns data points to clusters in a greedy manner, at each occasion allocating the unallocated data point that is closest to any of the prototypes (assigning that point to the prototype that is closest). After the assignment, the prototype of the is updated to reflect the new cluster member.

This clustering method (we will call it *prototype clustering*) is formalized in the pseudo code of algorithm 1. The set X is a set of data points (in this case a vector containing left and right forearm angles) to be clustered using the initial cluster centers (prototypes) c_i ($1 \leq i \leq K$). Furthermore, $X_i \subset X$ denotes the (initially empty) subset of X that belongs to cluster i, σ_i is the variance within cluster i (defined to be zero in case the cluster is empty), and $\overline{X_i}$ denotes the prototype of the set X_i, in this case the centroid of the vectors in X_i.

Algorithm 1: PROTOTYPE-CLUSTERING$(c_1, \cdots, c_K)$

$$\textbf{while } X \neq \emptyset$$
$$\textbf{do} \begin{cases} x, i \leftarrow argmin_{x,i} \frac{||x-c_i||}{1+\sigma_i} \\ X_i \leftarrow X_i \cup \{x\} \\ X \leftarrow X/\{x\} \\ c_i \leftarrow \overline{X_i} \end{cases}$$
$$\textbf{return } (c_1, \cdots, c_K, X_1, \cdots, X_k)$$

The selection of the data point x and and the cluster i to join depends on the variance σ_i of the cluster, in such a way that the agglomeration of data points into disperse clusters is easier than the agglomeration into compact clusters. In this way the criterion for adding a data point to a cluster becomes, informally speaking: 'how much is a data point inside the cloud of data points that form the cluster', rather than 'how far is the data point from the center of the cluster'.

Some other characteristics of this algorithm are that it allows for unequal population of clusters: Clusters do not necessarily get populated, depending on the structure of the data. Furthermore, in spite of its greedy nature, for a given data set, the outcome of the algorithm is robust to variations in initial prototypes.

We have run the *K-means* and the *prototype clustering* algorithms on the posture segments of the forearm angle data, using different values for the threshold parameter γ (eq. 1). The outcome of the various clusterings have been evaluated in terms of precision and recall with respect to the manual posture annotations.

Figure 8 summarizes the results, for the values of γ: 6, 10, 15, 20, 15, 20, 25, 30, 40, 50, and 60 degrees per second. Note that for both clustering algorithms, the threshold values that give optimal accuracies are roughly from 20 to 30 degrees per second. In this range *prototype clustering* slightly outperforms *K-means* clustering. Using $\gamma = 25°/s$, the precision and recall of both algorithms and for both conditions is shown in table 3. Figure 9 shows the corresponding cluster assignments of the angle-pairs graphically. Apart from the differences due to merging the clusters 1 and 2 on the one hand, and 7, 8, 9 and 0 on the other, the resulting clustering is similar to the manual annotations, as displayed in figure 5. Moreover, the effect of postures being more distinct during concert than rehearsal is also observed through the this clustering. This effect was confirmed using a Wilcoxon signed rank test ($z = -4.3493$, $p < 0.005$).

3.3.3 Analysis of transition speeds based on automatic annotation

Apart from the postures themselves, the transitions from one posture to another convey systematic differences between rehearsal and performance, as shown in section 3.2. In particular, songs where the forearm velocity during transitions is low on average (across rehearsal and concert), tend to show lower average forearm velocities during concert than during rehearsal. Conversely, in songs with higher

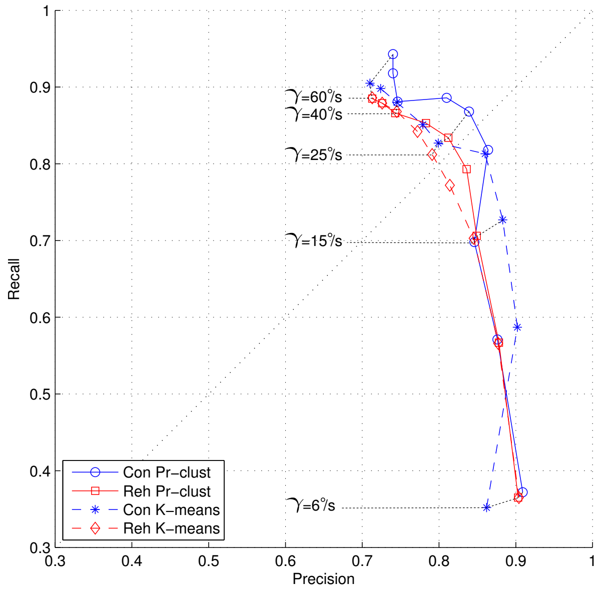

Figure 8. Precision and recall values for the discussed clustering method for concert and rehearsal data sets, using different values of γ (eq. 1)

	Rehearsal		Concert	
	Precision	Recall	Precision	Recall
K-Means	0.79	0.81	0.80	0.83
Pr-Clust	0.81	0.83	0.83	0.87

Table 3. The precision and recall of posture recognition using *K-means* and *prototype clustering* in the concert and rehearsal condition, respectively, using a threshold value of $\gamma = 25°/s$

average forearm velocities, forearm velocities during concert are higher than during concert (see figure 6).

Interestingly, this trend is also clearly visible in the automatic annotation, as shown in figure 10. The slope of the fitted line is 1.13, versus .64 when the analysis is based on manual annotation. Although the slope of the regression line is proportional to the threshold parameter γ (eq. 1), slopes higher than .66 were obtained for all settings of γ. This shows the persistence of the effect, independently of the parameters used for the automatic annotation.

4. DISCUSSION

The results presented in this paper are twofold. Firstly, in subsections 3.1 and 3.2, we have made an analysis of a singer's postures (and transitions) during rehearsal and concerts. This analysis is based on manual annotation of the data. Secondly, in subsection 3.3, we have proposed and evaluated a method to automate the annotation process. In this section, we will discuss the outcome of both parts, starting with the first.

The main findings presented in subsections 3.1 and 3.2 can be summarized as follows: Firstly, with an increase of almost 10%, a greater number of postures was registered during concert than during rehearsal. All postures observed during rehearsal were also observed during concert. Conversely however, three types of postures (8, 9, and 0) observed during concert, were not present during rehearsal. Thus, a wider variety of postures was used by the singer during concert. The open postures 5, 6, and 7 were more frequent during concert, at the expense of posture 1, that can be characterized as a closed position. Furthermore, during concert the transitions between two postures were over 20% longer on average than during rehearsal.

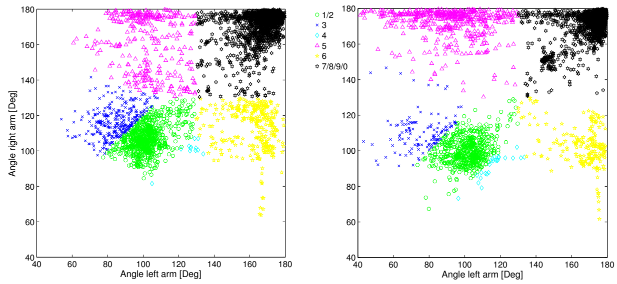

Figure 9. The angle of the left and right forearm extracted from accelerometer data for the rehearsal (left) and concert (right) condition. The different colors/symbols represent the different postures determined from automatic annotation using *prototype clustering* with $\gamma = 25°/s$

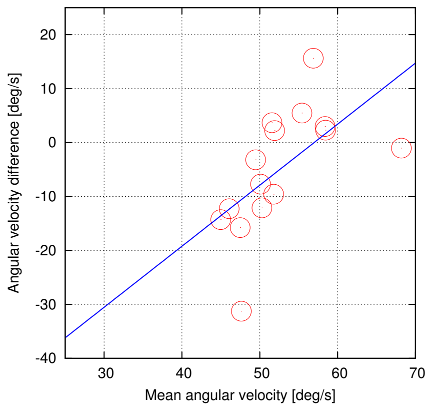

Figure 10. The correlation between the mean angular velocity and the difference in angular velocity for concert and rehearsal conditions. Each data point corresponds to the transitions occurring in a single song of the repertoire

The analysis of the motion sensing data revealed that in terms of left and right forearm angles, postures are less distinct during rehearsal than during concert, in the sense that the clusters representing the various postures (see figure 5) are more separated during concert than during rehearsal. An investigation of the mean angular velocity of transitions per song, as identified using the criterion in equation (1), showed that songs with low average angular velocity tended to have lower angular velocities during concert than during rehearsal, whereas songs with high average angular velocity tended to have higher average velocity during concert than during rehearsal (see figure 6). In other words, the way the singer moved between postures differs from song to song, and these differences are more pronounced during concert than during rehearsal.

From the above findings, it becomes clear that between rehearsal and concert, there are systematic differences in the singer's use of her body as part of the music performance. We assume here that the major difference between the concert and condition is the presence of an audience during the concert, and consequently a musical communication process between performers and audience which is absent during the rehearsal. Based on this assumption, a consistent view arises from the results, which is that the presence of the audience invites, or perhaps even requires the singer to use her body as a means of communication. It leads to a greater variety of postures that are more open in character. Furthermore, during concert the postures are more distinct in terms of the orientation of the forearms, making it easier for the audience to distinguish between different postures. This is in line with previous work on the role of the body in communication, which suggests that an open body position in contrast to a closed body position reinforces the communicator's intent to persuade [7, 8].

The systematic differences in angular velocity during transitions between postures from song to song, and the fact that these differences are amplified during the concert, strongly suggests that transitions between postures also have an expressive function. They may be used by the performer to emphasize changes or other significant moments of the music. The fact that transitions are substantially longer on average during concert can be partly explained by the larger distance between postures. But it can also indicate a stronger emphatic role of transitions during the concert. Further work is needed to investigate this possibility.

It is interesting to see these results in the light of the model of musical communication proposed in [9]. This model consists in the abstract view that the performer encodes his or her musical intentions in sonic and visual energy. This energy is received by the listener [1], who decodes the sonic and visual forms by a mirroring process in order to understand the performer's intentions, in a way similar in nature to the way other social behavior, like empathy, is thought to come about. Such behavior is thought to have a neurological basis [10], in the form of so-called *mirror neurons* (see [11]).

If we assume that the performer's musical intentions stay constant throughout rehearsal and performance, the fact that her corporal behavior is more articulate during the concert might be taken as an intent to encode the musical intentions in a clearer and more detailed way, in order to facilitate disambiguation by the listener in the decoding of the musical intentions.

The results presented in subsection 3.3 show that, with a relatively small amount of prior knowledge (viz. a description of the set of postures in terms of forearm angles), posture annotation can be done automatically based on motion sensing data, rather than manually using video recordings. The accuracy of the annotations in terms of precision and recall lies in the range of 80 to 90 percent. We have also shown that effects observed from manual annotations, such as amplification of differences in average angular velocity during transition can be reproduced using automatic annotations (see figure 6, and 10).

Lastly, a noteworthy result from the automatic clustering of postures based on angular data is that clustering accuracies are generally higher for the concert than for the rehearsal condition, as can be seen from figure 8. This is likely due to the fact that postures are more distinct during the concert, and it illustrates how the singer's efforts to communicate to an audience also facilitate recognition of postures using automatic methods.

[1] In this context 'listener' should be taken more generally to mean observer.

5. CONCLUSIONS

Although it is generally acknowledged that the presence of an audience has an effect on the various types of expression of music performers, not much is known about the way such changes manifest. With the aim of investigating the effect of an audience on the corporal expression of a singer in classical performance together with a gamba player, we have focused on the singer's postures and transitions between postures. Several systematic differences have been found between the dress rehearsal and the concert performances. The findings reinforce the view that the presence of an audience involves a musical communication process between performer and audience that leads to more articulate postures and movements, which are likely to improve the audience's understanding of the performer's musical intentions.

Furthermore, a method was proposed to automate the annotation of postures on which the analysis is based. Rather than manual posture annotation of video recordings, the automated annotation uses motion sensor data. The automatic annotation involves a novel data clustering technique, *prototype clustering*, that can accommodate prior knowledge in the form of cluster prototypes. This technique outperforms the *K-means* clustering algorithm initialized with the same cluster prototypes. An evaluation of this automated annotation method shows that it may be a viable alternative to manual annotation.

6. ACKNOWLEDGMENTS

This work was funded by the Flemish government, under the Methusalem-project *EmcoMetecca*.

7. REFERENCES

[1] A. Camurri, B. Mazzarino, M. Ricchetti, R. Timmers, and G. Volpe, "Multimodal analysis of expressive gesture in music and dance performances," in *Gesture Workshop*, pp. 20–39, 2003.

[2] J. Davidson, "Bodily communication in musical performance," in *Musical communication* (D. Miell, R. MacDonald, and D. J. Hargreaves, eds.), pp. 215–237, New York: Oxford University Press, 2005.

[3] R. I. Godøy and M. Leman, eds., *Musical Gestures: Sound, Movement, and Meaning*. Routledge, 2010.

[4] A. R. Jensenius, M. M. Wanderley, R. I. Godøy, and M. Leman, "Musical gestures: Concepts and methods in research," in *Musical Gestures: Sound, Movement, and Meaning* (R. I. Godøy and M. Leman, eds.), Routledge, 2010.

[5] S. Dahl, F. Bevilacqua, R. Bresin, M. Clayton, L. Leante, I. Poggi, and N. Rasamimanana, "Gestures in performance," in *Musical Gestures: Sound, Movement, and Meaning* (R. I. Godøy and M. Leman, eds.), Routledge, 2010.

[6] D. Moelants, C. F. Wu, M. Demey, and M. Leman, "Performing in concert and in rehearsal: a comparison using audio, video and movement data," *Proceedings of the 2009 European Society for the Cognitive Sciences of Music Conference (ESCOM)*, 2009.

[7] H. McGinley, R. LeFevre, and P. McGinley, "The influence of a communicator's body position on opinion change in others.," *Journal of Personality and Social Psychology*, vol. 31, no. 4, pp. 686–690, 1975.

[8] A. Mehrabian and J. Friar, "Encoding of attitude by a seated communicator via posture and position cues.," *Journal of Consulting and Clinical Psychology*, vol. 33, no. 3, pp. 330–336, 1969.

[9] M. Leman, *Embodied Music Cognition and Mediation Technology*. MIT Press, 2007.

[10] S. D. Preston and F. de Waal, "Empathy: Its ultimate and proximate bases," *Behavioral and Brain Sciences*, vol. 25, pp. 1–72, 2002.

[11] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi, "Premotor cortex and the recognition of motor actions," *Cognitive Brain Research*, vol. 3, pp. 131–141, 1996.