The Possibility of Digitally Recognizing a Person using Scent

Guus Frijters
University of Twente
Data Science (DS)
Supervisor: Prof. Dr. Ir. R.N.J. Veldhuis
Committee-members: Dr. Ir. L.J. Spreeuwers & Dr. C.G. Zeinstra

***Abstract*— In this study Gas Chromatography Mass Spectrograph (GC-MS) data was compared to look for similarities and differences, after which in future research a potential classifier can be designed. The data collection was done at Saxion University of Applied Sciences. Scent was collected from the hands of four persons which was then put through a GC-MS. The output data was used for the analysis using MATLAB. Difference analysis, cross correlations, threshold/peak detection and combinations of these three were done to compare the data. The amount of measured persons was too low to get to significant results. Based on the gathered data, it can be concluded that there are measurable differences between different persons. However, the cause of the differences remains unclear. The results of this study support the plausibility of discriminatory scent present in the human odor, however more analysis and test-subjects are needed to verify this hypothesis.**

I. INTRODUCTION

Scent recognition can play an important part in the juridical system. Bloodhound dogs are trained to match a certain scent to a suspect. These dogs were widely accepted as evidence in court [1]. In 2011 however, identification using the so called "geurproef" has been disregarded in the Netherlands due to the fact that it would be unscientific [2]. The argument for this was that the dog-trainer could steer the dog to a certain sample in order to get a verdict, which might be incorrect. Consequently, there is an urgency to investigate other, more scientific, ways to identify people using scent. Fortunately, it is already possible to measure body scent in more a scientific way. There is currently a five step process in which human body scent is measured using the germs present in the surroundings of the human body [4]. These germs are present in the bacteria which are constantly being shredded from the human body and picked up by a heat flow which travels in a heat-layer around the human body [5]. From this layer it is possible to extract the germs and put these through a gas chromatography (GC). The resulting chromatogram is converted to a digital form after which it is added to a database [4]. This method is currently only used to create a database and is therefore not yet used for recognition.

Unfortunately, all GC's are different due to the carrier gas, temperature and column that is used. Next to that, a person's scent is highly dependent on several factors. Therefore it is important to get insight in any discriminatory supplements in a person's scent and, if so, find out if it is possible to make a classifier for these supplements. With a correct classifier, scent based evidence is not dependant on bloodhounds and their trainers anymore and can therefore be used in court again. In order for this to work, collecting and measuring scent-samples also need to become easier. To make this possible, it is first needed to look at a person's scent on its own and the possibility to analyse it. This results in the question whether it is possible to filter out the discriminatory information from a person's scent and if it is possible to define a classification for it. Therefore, the research question was formulated: Is it possible to distinguish a person using their scent?

II. THEORY

A. Person's scent

Person's scent consists out of 3 layers [4]. The first layer is influenced by a person's environment. For example the places he/she lives and works in, that are detectable in our scent. The second scent layer is influenced by a person's diet. These two layers are not constant over time. The environmental layer can change quite drastically while the diet-layer is semi-consistent over a period of time [4]. The third layer is a person's personal scent. Since this is person bound, this layer is a more constant over time. A person's scent is related to the ABCC11 gene, which could indicate that there is some discriminative evidence present in scent [8]. To get more accurate results with scent identification, research should focus on the personal third layer of human scent, and its discriminatory information. Therefore, it is important to filter out the two non-discriminatory layers.

B. Existing methods

Scent recognition is already an upcoming biometrical method. Forensic analysts are researching possible methods to have a digital scent-recognition [3], [4], [5], [6], [9].

Like a fingerprint and DNA, the discriminative third layer of human scent can be used as an identifier for persons. This is already done in a study where 13 subjects where used and an average recognition rating of 85% was achieved [3]. This proves that is possible to identify humans, based on their discriminatory information in their body odor.

Various researches shows us different methods to analyse a person's scent [3]. To discriminate body oder, five types of classifiers were tested; K-NN (Nearest Neighbors), LDA (Linear Discriminant Analysis), Logistic Regression, Naive Bayes and linear Support Vector machines (SVMs) [3].

A different method to analyse mass spectrometry data is the use of peak detection. Research tried to find the best algorithm possible for Peak detection [6]. However, peak detection is not directly a comparison method but can be used to compare the peak-locations of different samples. Peak detection is dependant on the intensity of the peaks and the locations, and when these are known, a different method is still needed to compare these locations.

In other research Spearman Rank Correlations are used to determine overlap between different scent samples [5]. Multiple subjects were measured multiple times and were paired using Spearman correlation. This was also done after removing known substances from the dataset in order to get a ’primary scent’ [5].

It is also possible to use different filtering approaches to analyse body odor [9]. These consist of a CFS filter, Linear correlation filter, Rank Correlation filter (as mentioned before), Relief filter and Principal Component Analysis (PCA). PCA, linear correlation filter and LDA are closely related in how they operate.

Another method is to identify all the different possible substances present in scent and use pattern recognition to search for these different substances [4]. This means that a big data set is needed to measure all the different substances.

These methods make use of a large amount of data available to compare and are then improved based on the results. They are all used in the aspect of machine learning and are mainly used for research-purposes and database creation.

III. METHODS

A. Data collection

The data collection for this research was done at Saxion School for Applied Sciences. Scent samples were captured from the participants’ hand palms in three ways. The collected scent was put through a Gas Chromatography Mass Spectrography (GC-MS). These measurements resulted in a chromatogram, which were converted to digital data sets.

At the start of this research there were two measurements available from previous research. Scent samples were captured from a hand palm of a participant using a plastic sandwich bag in which the hand was held for 30 minutes. After the 30 minutes, the air inside the bag was put through a GC-MS. These measurements consisted of 3 measurements of person A and 4 measurements of person B. These measurements verified that it was possible to measure scent.

In addition more recently, the scent of two more persons were captured (person C and person D) in two different ways. The first way of measuring consisted of a bag in which the hand is held for 40 minutes. During this time, the bag would act as a steam room and generate a sample.

This sample contained the scent of the hand that was inside the bag as well as potential waste. To know which part of the measurement was from the hand and which part was waste, a zero measurement was done. This zero-measurement measured the ”scent” of the bag. This means that the zero-measurement could be subtracted from the person measurement and this would result in a graph wherein only the persons ’scent’ is present.

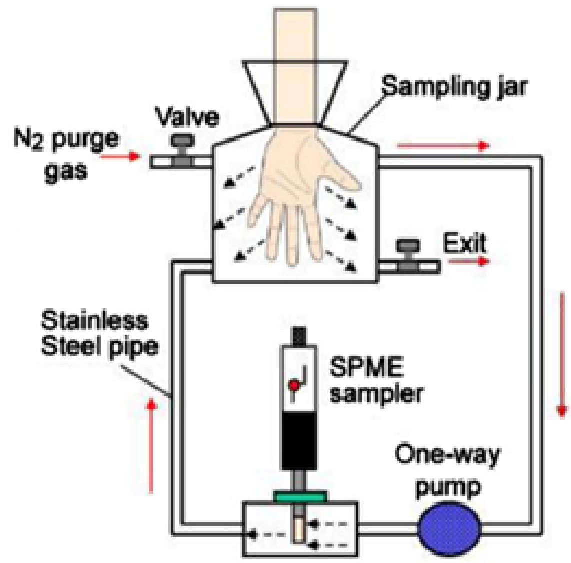


Fig. 1: Schematical view of the measurement

The second way of measurement was done with the use of a small tube. This tube was hollow and had one open end which was connected to the GC-MS. This tube had small holes. This tube was held for about 40 minutes by a person. The tube than acted as a sort of steam room in which the air contained a lot of person-specific scent. With this measurement a zero-measurement was also needed to subtract the tubing from the final measurement.

In both setups, the scent was captured in a sampling jar (Figure 1). In the first method the bag acted as a sampling jar and with the second method the tube was the sampling jar. The Solid Phase Micro Extractor (SPME) extracts the sample from the sampling jar and injects it in the GC-MS.

Person C was measured twice with both methods, while person D was measured once with the bag-method and twice with the tube method.

The newer measurements were more elaborate than the earlier ones, these were done in duplicate, and a correct zero measurement was taken. This also meant that more analysis was possible.

B. Data Analysis

Since the data collection resulted in a sample set of seven good samples, analysis had to be simplified in the following methods.

Before the data analysis could be done, the zero-measurements done at Saxion University for Applied Sciences were subtracted from the measured data where possible. After the zero-measurements were subtracted and the negatives were removed. The data was analyzed in three different ways to see if it was possible to distinguish the different persons from each other.

In the first analysis, a binary sample set was made. The data from person C and D was converted in high’s (ones) and low's (zeros) to gain insight in the peak pattern that was present in the sample. These peaks could then be compared to another sample set. This was done with different thresholds (intensity of 10000, 20000 and 30000) to see the different effects.

Another way the data was compared was differential analysis. The duplicate measurements were subtracted from one another. If the samples would be the same, the resulting product would be zero.

The third analysis was to apply cross correlation to look for similarities in the different data sets. With a cross-correlation it was possible to see how much two functions (or samples) match with each other. Two samples were multiplied to get their power. If the samples were identical, the power was high when the functions were exactly on top of each other. To make this more visible, the normalized cross correlation were taken. This means taking the cross correlation, and dividing by the normalized functions of themselves. This results in a spectrum from 0 to 1, where 1 means a perfect overlap. Different correlations were taken. Firstly, correlations where a person's sample was correlated with the duplicate of the person were done. Secondly, correlations were done where the different methods (bag and tube) of the same person were correlated. Next to that, also correlations were taken with different persons (person C correlated with person D) to see how much overlap there was between persons.

The results of these methods indicated that it was useful to add certain analyses. It was clear that it could have effect to remove the last part of the measurement, which was caused due to the measurement setup. Therefore the last fifteen minutes were removed and not used in further analysis.

Without the last fifteen minutes new analyses took place. The first of which was cross-correlations in the same way as before (Person with it's duplicate, the different methods and the different persons). Also correlations were done after threshold-analysis (with the lowest threshold). This would show whether the location of the peaks matched.

Next to that, correlations took place where the samples were partitioned per five minutes. These analyses were done to gain insight in the correspondence between the duplicates at different times during the complete measurement.

IV. RESULTS

A. Data Collection

The first measurement method resulted in the data from person A and person B (Figure 15a and 15b in the appendix). The second bag method resulted in the data from person C and D (Figure 15c and 15d in the appendix). The tube method resulted in a second sample set of person C and D (Figure 15e and 15f in the appendix) These results show that persons A and B are different in many aspects (amount of peaks and height) than persons C and D.

B. Data Analysis

Subtracting the zero-measurements (Figure 16a and 16b in the appendix) and canceling out all the values below zero resulted in person C (Figure 2) and person D (Figure 3). There were no zero-measurements from person A and B which meant that it was not possible to subtract the zero-measurement. Therefore no further analysis was done with the measurements of person A and B.

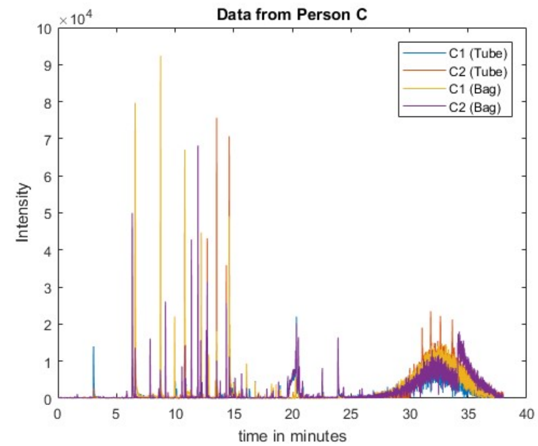


Fig. 2: Corrected data from Person C

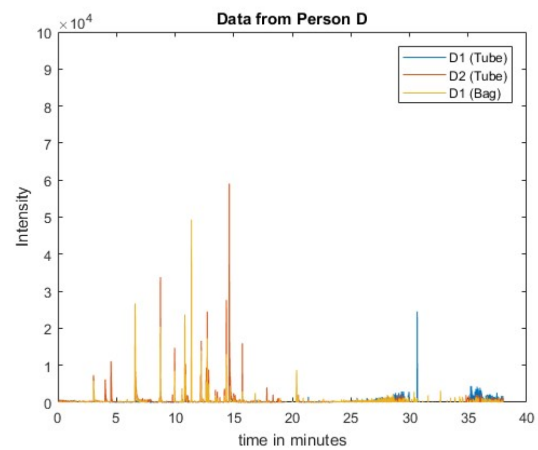


Fig. 3: Corrected data from Person D

A threshold comparison is done in order to look for differences in peaks and intensity between the different measurements. The results can be seen in Figure 4 and Figure 5. It is clear that this analysis is highly dependent on what threshold is taken. Lower thresholds result in more high's (ones). The figures show that both person C as well as person D have a high amount of peaks around the 10 minute mark. It can also be seen that person D has lesser peaks above the thresholds than person C. It can also be seen that with a higher threshold, less peaks are visible.

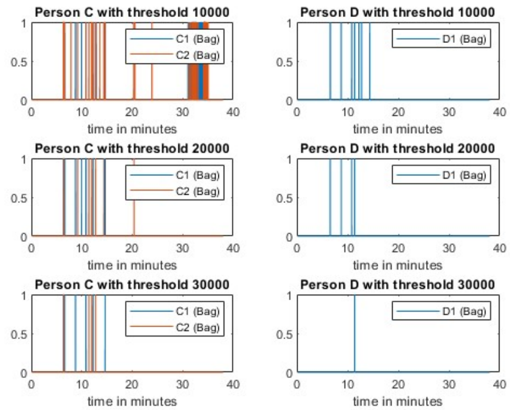


Fig. 4: Different thresholds of the Bag measurements

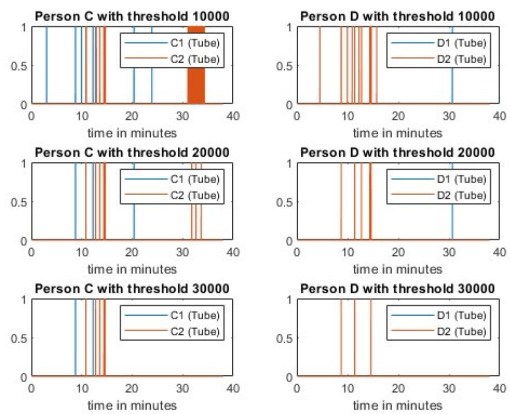
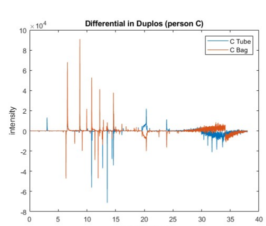
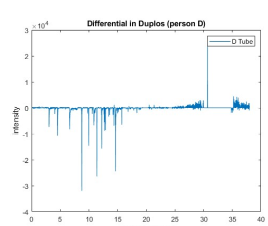


Fig. 5: Different thresholds of the Tube measurements

Differential analysis shows that there are big differences between the duplicate measurements (Figure 6a and 6b). This shows that the difference within one person is still high, which would indicate that the measurements did not measure the same person at all. Different persons and different methods showed also a big amount of differences (Figures 17a and 17b in the appendix).



(a) Person C



(b) Person D

Fig. 6: Difference analyis

The normalized cross correlations resulted in Figure 7 where all the duplicate-measurements are correlated. This shows that the duplicates are similar with each other. The different methods showed more differences (Figure 8). It is interesting to see that the first and second method seem to be more influential with person D than with person C since the normalized cross-correlation is higher than a half (which it also is with the duplicate). From both the duplicate as well as the different methods, there is a high amount of noise present in the measurements of person C. Correlations between persons showed that the similarities between different persons were not that high (Figure 9).

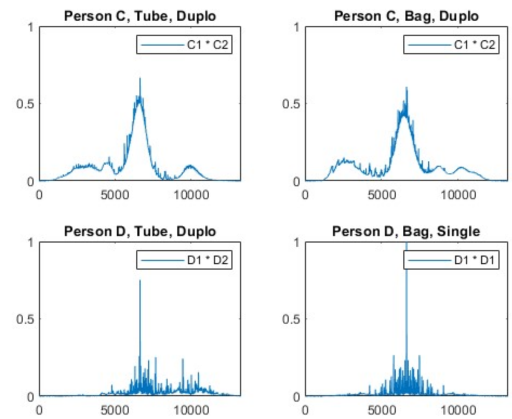


Fig. 7: Normalized Cross Correlation of the Samples

The removal of the last fifteen minutes of data resulted in Figures 18a, 18b, 18c and 18d in the appendix. The cross-correlations done after the removal of the last fifteen minutes showed that the duplicate measurements for person C did not match as much as with the last 15 minutes (Figure 10).

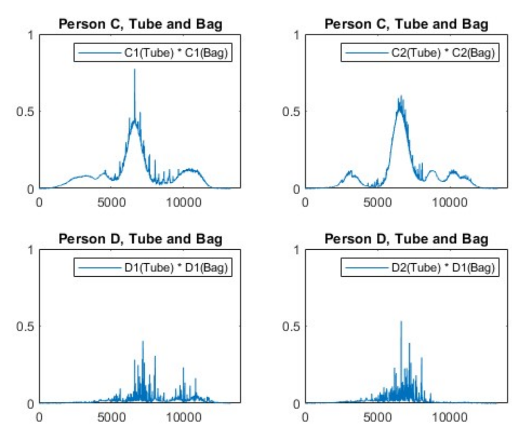


Fig. 8: Normalized Cross Correlation with the different methods

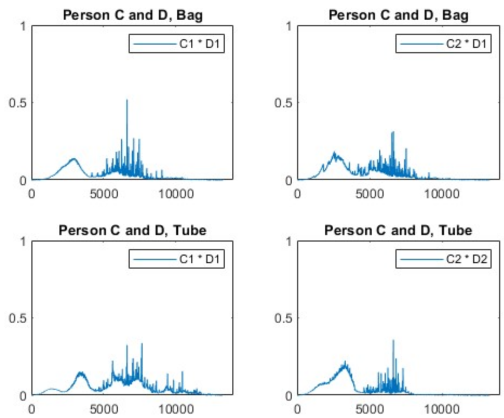


Fig. 9: Normalized Cross Correlation with the different Persons

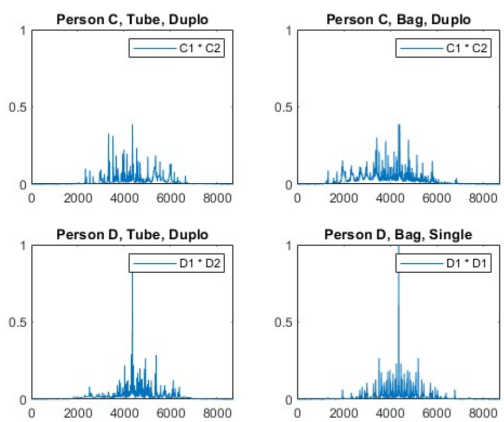


Fig. 10: Normalized Cross Correlation of the samples without the last 15 minutes

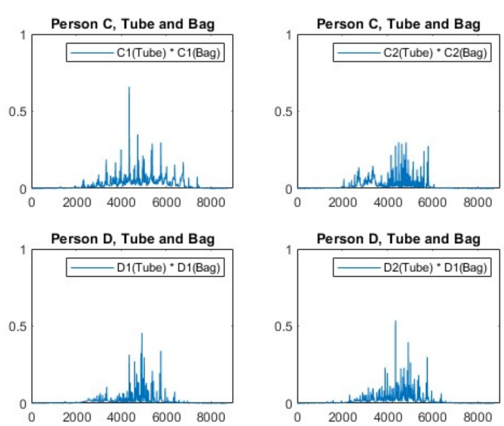


Fig. 11: Normalized Cross Correlation with the different methods without the last 15 minutes

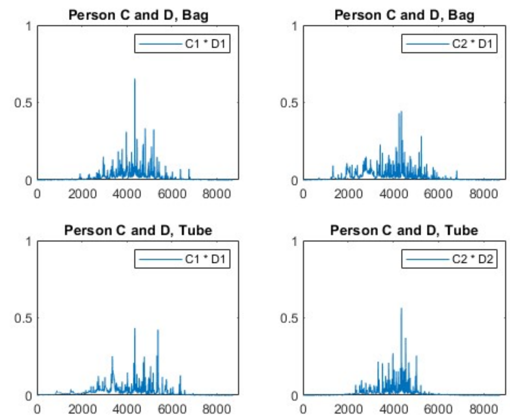


Fig. 12: Normalized Cross Correlation with the different Persons without the last 15 minutes

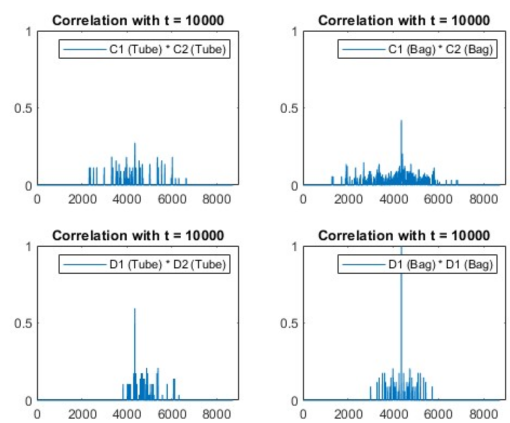


Fig. 13: Normalized Cross Correlation after thresholds

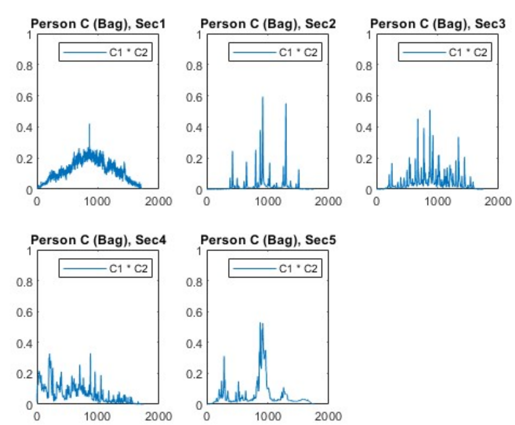


Fig. 14: Normalized Cross Correlation with different sections

The correlations between the different methods (Figure 11) without the last fifteen minutes shows that person C is less dependent on what method is used than person D. The correlations with different persons show that they have quite some similarities, with the first bag measurement of person C and D a high value of 0.65.

Cross-correlating after a threshold analysis resulted in Figure 13. This shows how the peaks correlate. It can be seen that the peaks of person D correlate better than the peaks of person C. The tube measurement of person C even has a maximum correlation of only 0.27.

Partitioning the samples resulted in Figures 19a to 20c in the appendix. This is a partitioning per 5 minutes. With the partitioned samples cross-correlation was done for each part. This resulted in Figures 14, and Figures 21a and 21b in the appendix. These Figures show that at certain parts of the measurement, the samples show more overlap than in other parts.

V. DISCUSSION

One of the issues that occurred during this research was the data-collection. The data collection had some hiccups which caused a delay in the creation of a sufficient data set. Therefore it was not possible to do a full data analysis. Since there were only two persons measured, it was challenging to compare different people with each other.

Next to that, at the start of this research, it was decided to measure 10 more persons in this manner in order to have a larger set of subjects to work with. During the research it turned out that it took a lot of time to gain sufficient data to conclude something about the data. In the beginning it only consisted of two persons (person A and person B) which where used to get familiar with different kind of methods to evaluate the data. When new data was send, it turned out that the data collection before this research (person A and person B) was not correct. The data from person C and D showed a big difference in the amount of peaks. Saxion University of Applied Sciences figured out that the cause for this was that the earlier measurements were done incorrectly. It showed that it was possible to measure scent, however this scent most likely did not only consist of human scent, but also a lot of waste substances, which made it clear that they were not comparable. Since also no zero measurements were done with person A and B. No further analysis could be done with these measurement. Next to this, the results of person A and person B had a different amount of samples. This made comparisons in any way difficult to do in MATLAB. It was first needed to interpolate the data or remove part of the data in order to get and equal amount of samples. Researching the best way of doing this took time, and was not needed with the newer measurements.

Another point of discussion is the gas chromatography in itself. A GC is useful when comparing samples directly. However, since there is as of this moment no standard GC, it is difficult to compare data from different GC's. This is due to the fact that the results of a GC are dependent on the specifications of the GC. Which carrier component is used, the temperature of the measurement and the intake method can all be varied. This means that if there is data from different moments, different GC's and/or different methods, it is not fair to compare these data sets. This also implies that there are certain standardized methods needed before scent recognition can be used as a standard recognition method.

Another problem is that there is no normalized intensity in the measurements. The intensity is only an indication in amount, and is therefore unitless. This means that it is difficult to decide on a correct threshold to compare the different samples since this will not be based on any real evidence.

The same zeros and ones in the threshold measurement would mean that the sample could be from the same person. However, it could also be the case that two persons wear the same deodorant which would give enough similarities that this could indicate that these two samples are from the same person. For this analyses an estimation for a good threshold is needed. However, the data made it clear that such a threshold does not yet exist.

The differential analysis also had some points of discussion. Interpolating causes issues with GC-MS results since all peaks can mean different substances. Also the absence of peaks have meaning. However, the newer measurements were done more properly and all the data was of equal length. Therefore interpolation was not needed anymore.

A different point of discussion occurred with the cross correlations. The hump at the 30 minute point was present in all the measurements, so this likely was caused by the substances the GC which was measured. Therefore all the cross-correlations showed a lot of similarities. This was confirmed by Saxion University of Applied Sciences. Therefore these last 15 minutes were removed for the later part of analysis.

The fact that person C did not correlate well with itself, could mean that the different layers of scent were present. This could be made visible with a partitioning of the data set to show if it is similar at certain timestamps. This could indicate that important discriminative information is visible at a certain point in time.

With a partitioned threshold analysis, there are two possible ways to set a section as high (or low). Both are unfair since with a GC-sample it is important at what point a peak is placed in order to identify a substance. Two peaks in one section can mean two completely different things. Therefore partitioned threshold analysis can show some match, but cannot be used as definitive.

VI. CONCLUSIONS

From the results in this research, it seems to be possible to digitally make a distinction between different persons using their scent. However, it has only been shown with two persons at this point.

From the threshold measurement it can be concluded that comparison of persons scent is too dependent on a chosen threshold. If the peaks were roughly the same we can conclude that there are enough substances in equal amounts present, which could indicate that it is from one person.

From the differential analysis can be concluded that the persons are not even slightly similar. However, this can also mean that only the intensity is too different from each other or that the different layers are present.

With the cross correlations, the data shows that there are differences in the measured persons. Person C is different from person D. However, person C is not consistent with itself at some points.

The last fifteen minutes of the samples created a lot of overlap between different persons and methods. This shows that the method is at this point not the best.

This research shows that it is possible to visualize differences between persons and that they therefore are distinguishable.

VII. RECOMMENDATIONS

From this research it seems to be possible to recognize persons using their scent. However, there are still a lot of if's and uncertainties. Therefore there are some recommendations for further research.

First of all, it is important to have more measurements with the same method. With this research, only two persons where measured twice with the same method. With this data it was possible to make some assumptions about human scent, but it not possible to make any definitive conclusions. It is therefore important to get more (sufficient) data for further research.

To gain more insight in the human scent or to make identification easier it might be useful to sample the substances we would like to filter in the human scent. If we for example want to know if person X is a man or a woman, it would be useful to know which part of the data-sample is testosterone or estrone. This can be done by creating a sample of testosterone and measuring it with the GC-MS which will be used to sample all the subjects. If the peaks of testosterone or estrone are present in the data of the subject, we could easily say something about the sex of the subject.

Another point for future research is to have a better look at the zero-measurement. In this research the last fifteen minutes were discarded because it is thought that this is waste in the measurement from the measurement-setup. It should be the case that this would be removed using the zero-measurement. It therefore might be a good idea to scale the zero-measurement in such a way that the hump from the setup is completely removed.

It could also be useful to take correlations on the separate parts of the data. For example, cut the sample up in pieces of 5 minutes (or less) and correlate these samples. The same can be done with differential and threshold analysis. This has been partly done, but more research is needed with different subsets.

Next to this, it is also a good idea to look at existing methods and repeating some of their methods on the received data. This might give insight in the received data. More data is needed for most of these methods.

To achieve better results, it would be a good idea to measure persons over longer period of time (for example, one year long, every week). This could make the different layers visible. If this would be combined with for example a partitioned analysis, it might give insight in what parts of the GC-MS data are the discriminative parts.

ACKNOWLEDGMENT

I like to thank my supervisors for their insights and their feedback on te process. I would also like to thank Reinier Groot for his his data collection at Saxion University for Applied Sciences.

REFERENCES

[1] T. Tomaszewski, Scent identification evidence in jurisdiction (drawing on the example of judicial practice in Poland), Forensic Science International 162 (2006)

[2] H. van Maanen, Geurproef dient niet meer als bewijs in strafzaak ,Volkskrant 22-04-2011

[3] I. Rodriguez-Lujan, Analysis of pattern recognition and dimensionality reduction techniques for odor biometrics, Knowledge-Based Systems 52 (2013)

[4] P. Inbavalli, Body Odor as a Biometric Authentication, International Journal of Computer Science and Information Technologies, Vol 5 (2014)

[5] A. M. Curran, The Differentiation of the Volatile Organic Signatures of Individuals Through SPME-GC/MS of Characteristic Human Scent Compounds, Journal of Forensic Sciences Vol. 55 No.1, (2010)

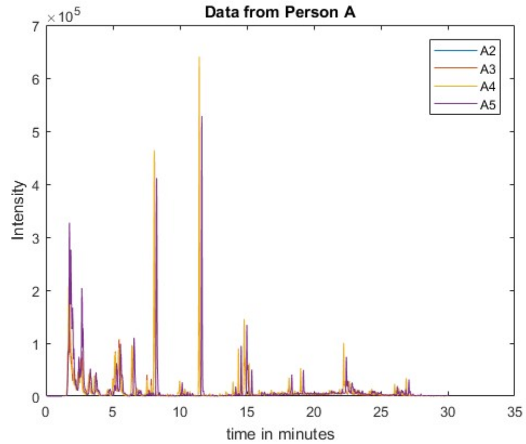
[6] Chao Yang, Comparison of public peak detection algorithms for MALDI mass spectromety data analysis, BMC Bioinformatics 10:4, (2009)

[7] Anil K. Jain, An Introduction to Biometric Recognition, IEEE Transaction on Circuits and Systems for Video Technology, Vol. 14 No. 1 (2004)

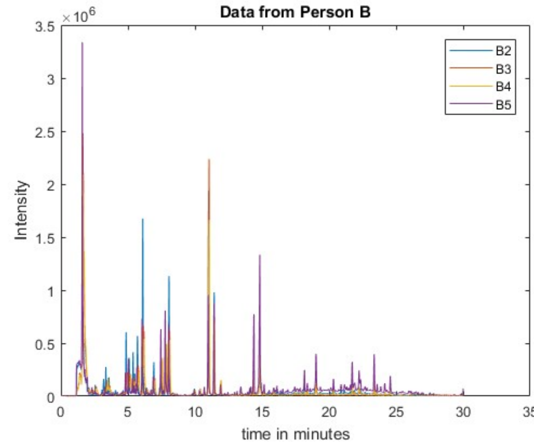
[8] A. Martin, A functional ABCC11 allele is essential in the biochemical formation of human axillary odor, Journal of Investigative Dermatology 130 (2010)

[9] S. Kr. Jha, Body odor classification by selecting optimal peaks of chemical compounds in GC-MS spectra using filtering approaches, International Journal of Mass Spectrometry 415 (2017)

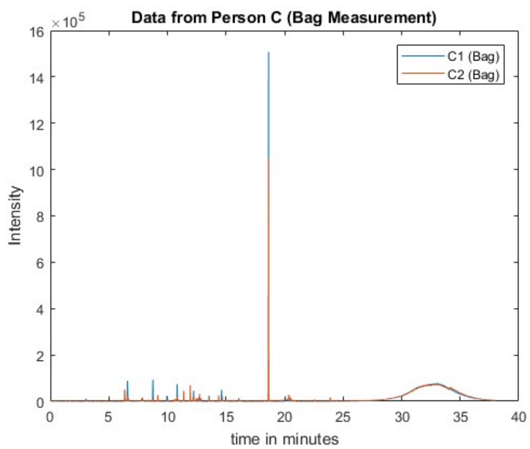
APPENDIX



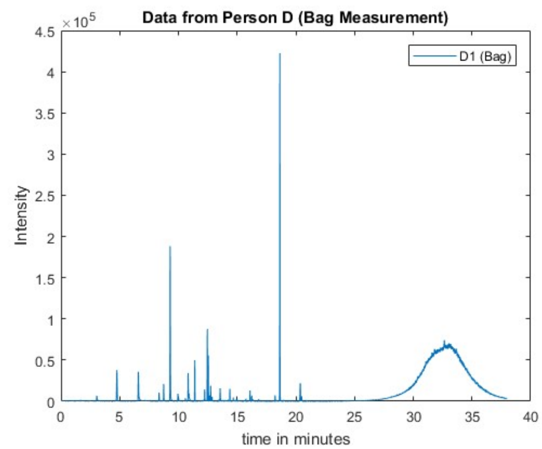
(a) Raw data from Person A



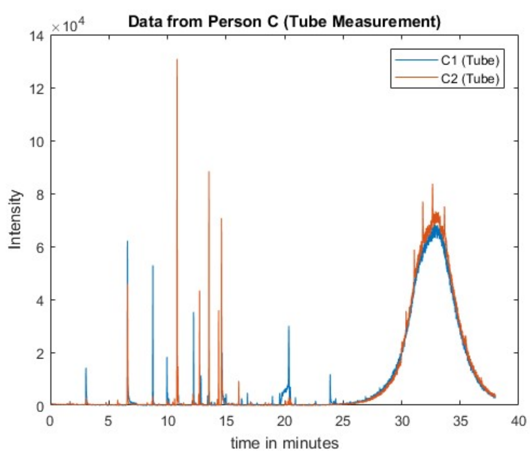
(b) Raw data from Person B



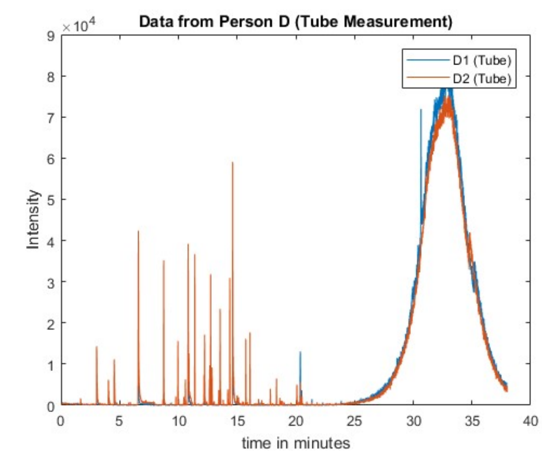
(c) Raw data from Person C (Bag)



(d) Raw data from Person D (Bag)

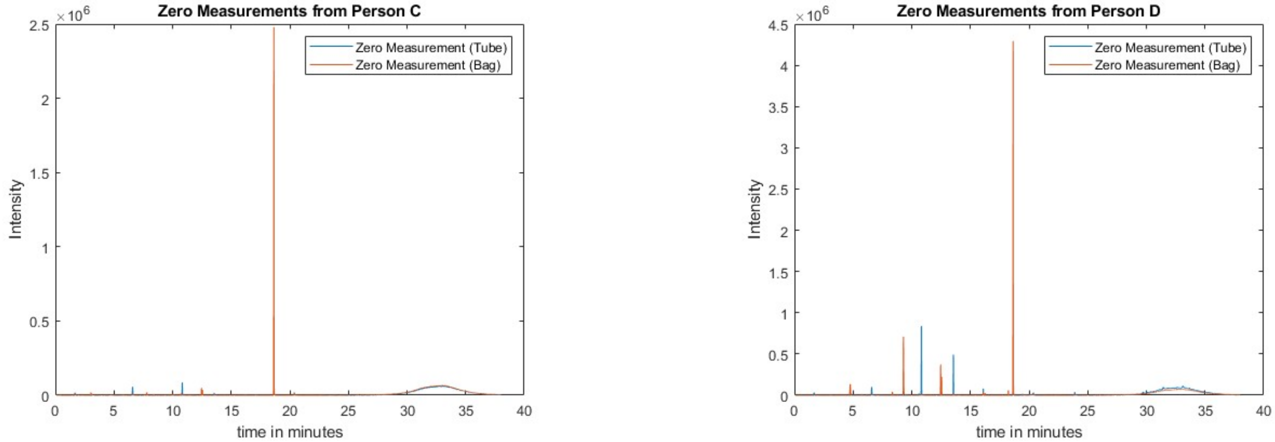


(e) Raw data from Person C (Tube)



(f) Raw data from Person D (Tube)

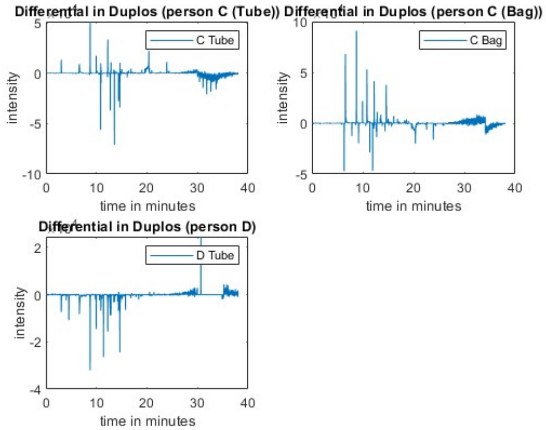
Fig. 15: Raw data



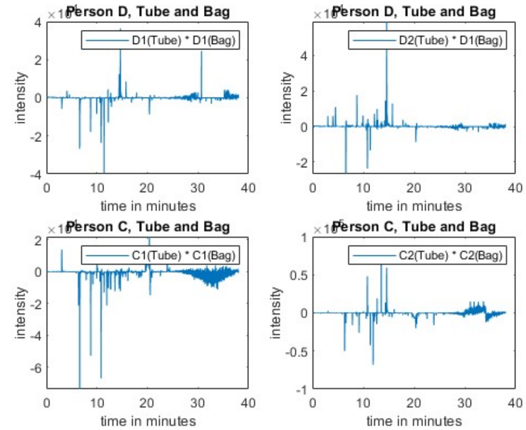
(a) Zero Measurements from Person C

(b) Zero Measurements from Person D

Fig. 16: Zero Measurements

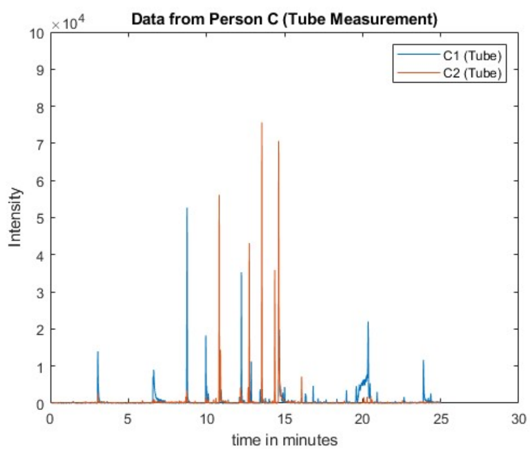


(a) Difference between Persons

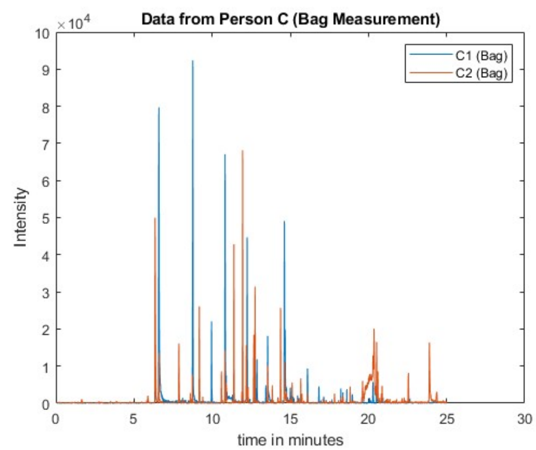


(b) Difference between methods

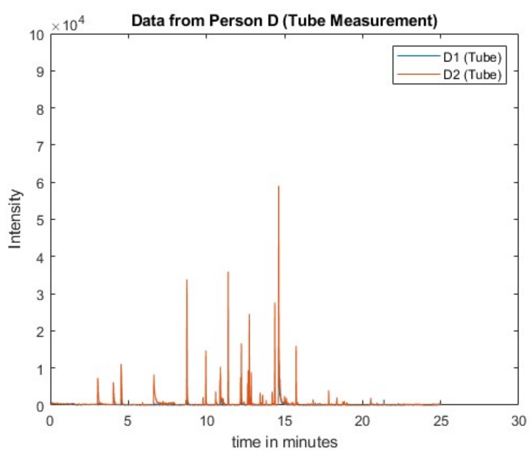
Fig. 17: Differences



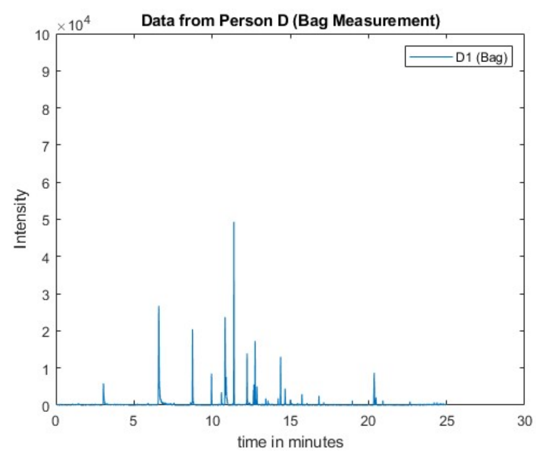
(a) Data from Person C (Tube) after removal of last 15 minutes



(b) Data from Person C (Bag) after removal of last 15 minutes

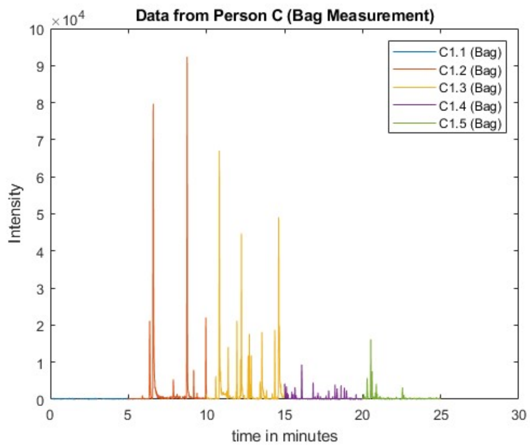


(c) Data from Person D (Tube) after removal of last 15 minutes

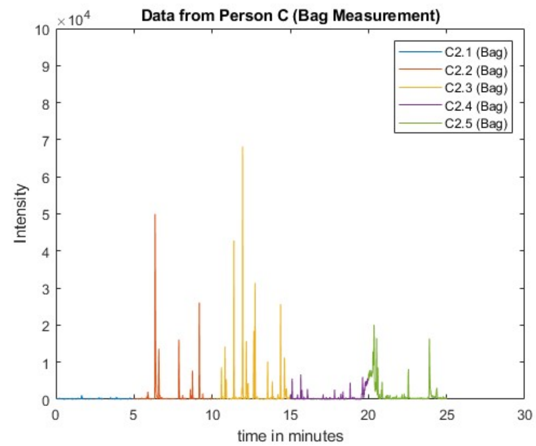


(d) Data from Person D (Bag) after removal of last 15 minutes

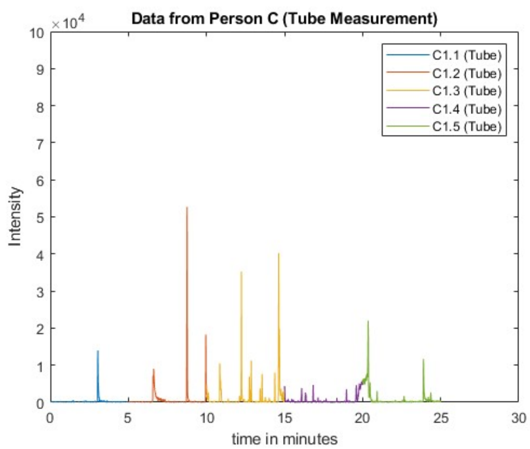
Fig. 18: 15 minute removal



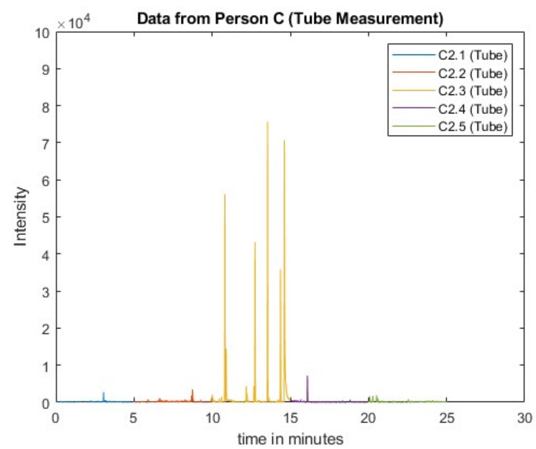
(a) Data from Person C1 (Bag) after sectioning



(b) Data from Person C2 (Bag) after sectioning

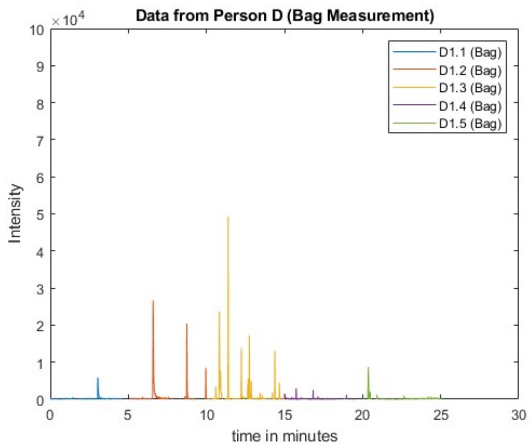


(c) Data from Person C1 (Tube) after sectioning

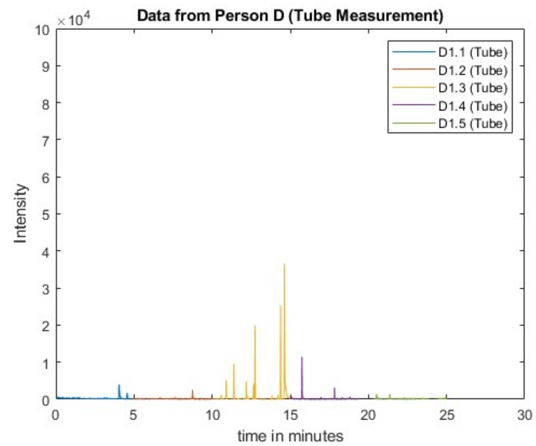


(d) Data from Person C2 (Tube) after sectioning

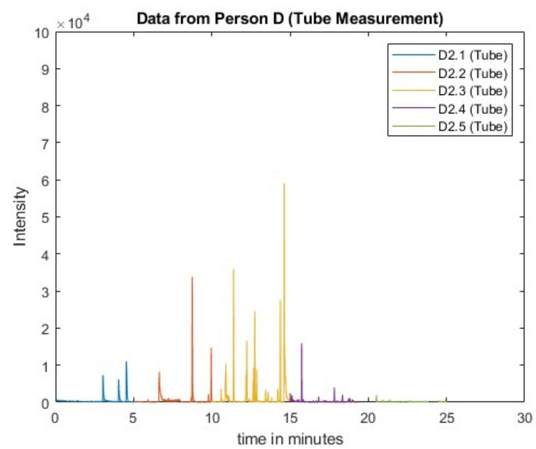
Fig. 19: Sectioning of person C



(a) Data from Person D1 (Bag) after sectioning

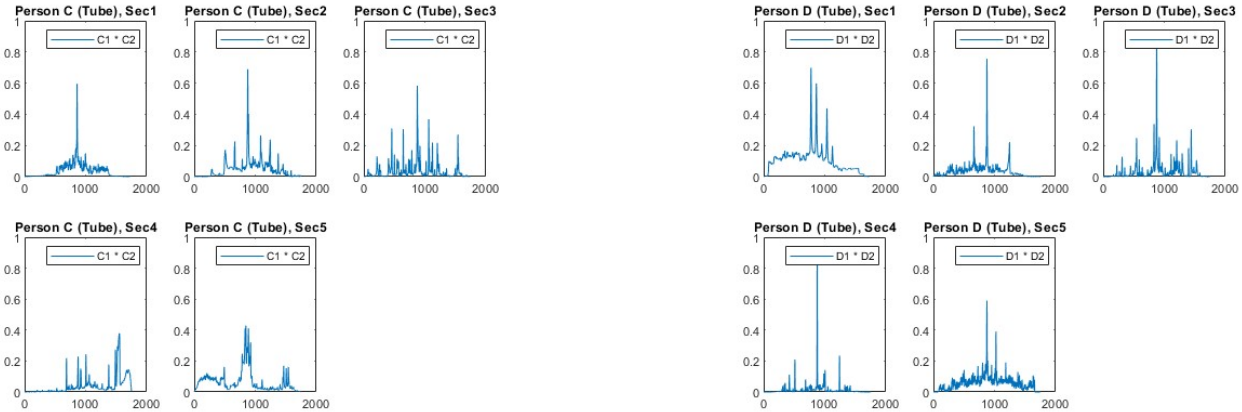


(b) Data from Person D1 (Tube) after sectioning



(c) Data from Person D2 (Tube) after sectioning

Fig. 20: Sectioning of person D



(a) Person C, Tube

(b) Person D, Tube

Fig. 21: Normalized Cross Correlations with Different Sections