Course Work Cover Sheet - the School of Computing

Your name: Kyle Harrison

Matriculation number: 11000 9870

Module code:AC42001

Assignment title (see module guide): Lab Report 2

Assignment number (see module guide):2

Date submitted: 3/02/2014

Date due to be submitted: 3/02/2014

Number of hours you spent on this item of coursework:

Numbers of hours you were advised to spend (see module guide):

Word count (if applicable): 2,800

Your course work may be used by the School of Computing for demonstration purposes with other classes and/or members of the public, unless you indicate otherwise.

Tick the box at the right if you do *not* wish your work to be used in this way.

I certify that this assignment represents all my own work and that no parts of it have been copied and that no collusion has taken place with any other person. I have also read and understood the School of Computing's regulations as regards course work and I certify that this course work was performed according to those regulations.

I certify that any files submitted with this assignment have been virus checked and have no associated viruses.

Signed: Kyle harrison

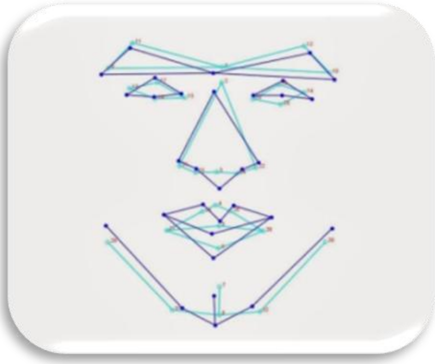
Lab Report 2

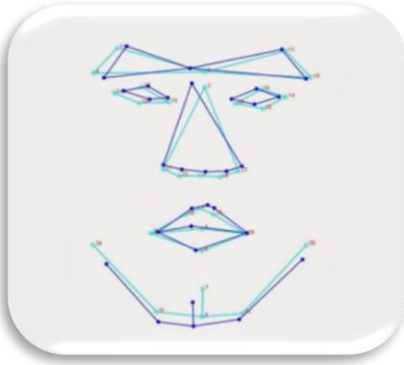
Acquiring and Land marking a Face Data set

Problems with land marking the data became apartment when trying to read TPS files. Problems with missing IDs for each set that needed to be appended to the list was quite common and often simple to fix when comparing the image title against the images on the shared drive however problems with land marking were not so easily correct. With a small dataset it was easy but time consuming to change some IDS however for a much larger dataset this would not be feasible.

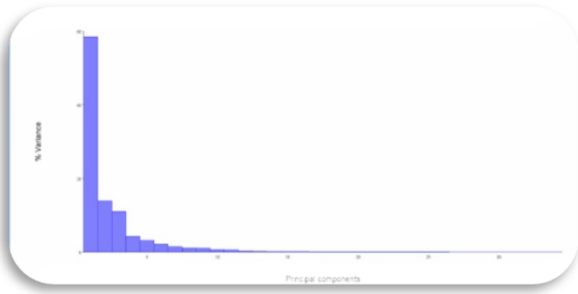
It is likely these problems can be attributed to a lack of experience as frequently I discovered the addition or subtraction of a land mark or even land marks placed in the wrong order resulted in out liars and the image needing to be recreated. From the photos collected, a lot of images seem to be captured at a tilt possibly demonstrating how scale, transformation, and rotation is carried out within MorphoJ however the main problem may have come from land marking images where individuals had beards, long hair, glasses, etc.

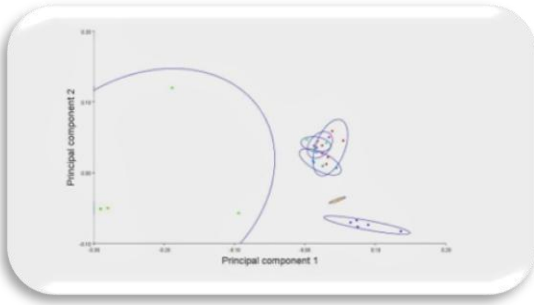
Using a small subset of the data the wireframes for Principal components 1 & 2 can be seen below. These results have been heavily skewed by out liars in the data set.



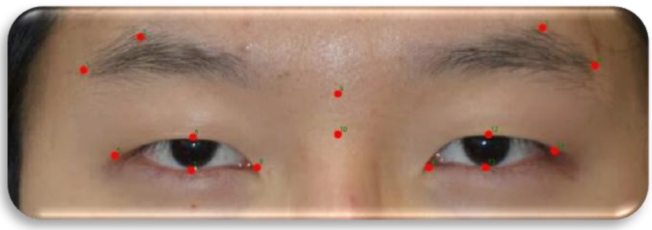


PC1 was over 60% of the variance and identified something in the neutral faces was not correct.





The Scatter plot identifies that one individual has the largest variance in PC1 and PC2 compared to all other individuals. Colour coding and setting confidence eclipses it can be seen that the maximum variance is in that individuals neutral face which would be expected to have the least. Reviewing the TPS file shows that the landmarks are incorrectly identified.

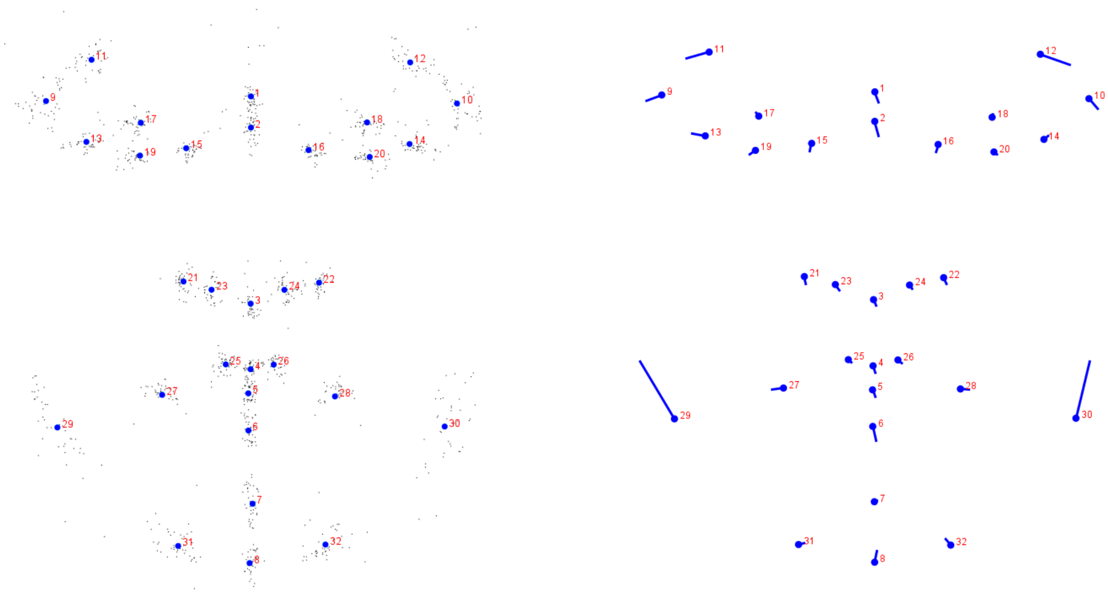


In the wireframe above this caused features to become skewed and malformed and altered the PC to a very large degree. Removing .TPS files like this allowed for a normalised dataset to be produced.

Analysis of Neutral Faces

Principal components analysis of the shapes in all images of the class (with the exception of my own) in which they exhibited neutral facial expressions has been carried out with the exclusion of individuals that have been identified as out liars.

Lollipop Graph



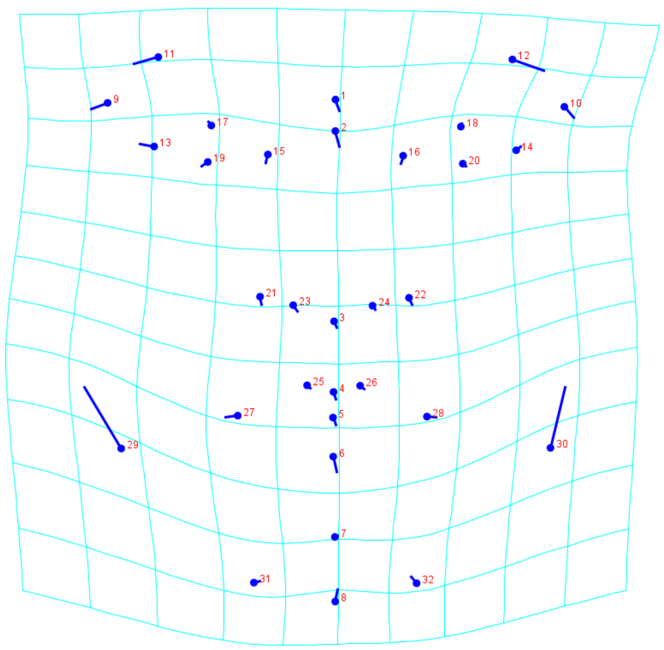
The lollipop graph shows that Principal component one is the Protrusion of the mental tubercle on both sides becoming elongated meanwhile the brow line extends slightly upwards whilst the other facial features remain unchanged. The mandibular corpus can be used to determine the approximate age of the individual and jaw movement and the shape of the mandibular corpus can be used to used to link to dietary patterns.

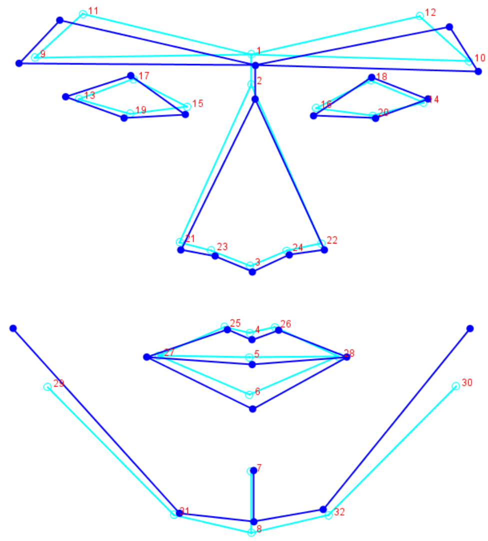
"In experiments on modern human subjects, [Agrawal et al. (2000)]() found that the mastication of hard foods (foods of low *R*/*E*) was associated with larger lateral excursions of the mandible than was the case with softer foods. Such hard foods are exactly those that would be predicted above to lead to thicker enamel in order to resist deep (radial) EDJ-type cracks. To the extent that jaw structures might be linked to such dietary patterns, then it is possible that a hard diet requiring wider excursions might lead to the development of a relatively broader mandibular corpus."[1]

With individuals from very distant geographical regions of the world and a high likely hood of different dietary patterns between the two groups it could be inferred that the variance in PC1 can be used to identify individuals from different geographical locations.

[1] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2409106/

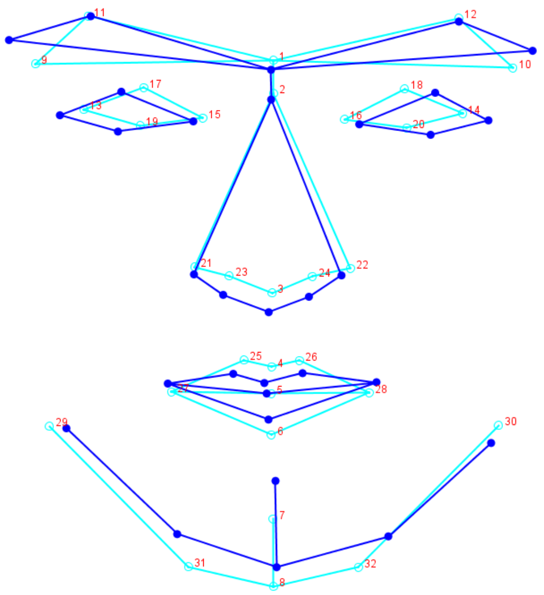
Transformation Grid and Neutral Face Wireframe



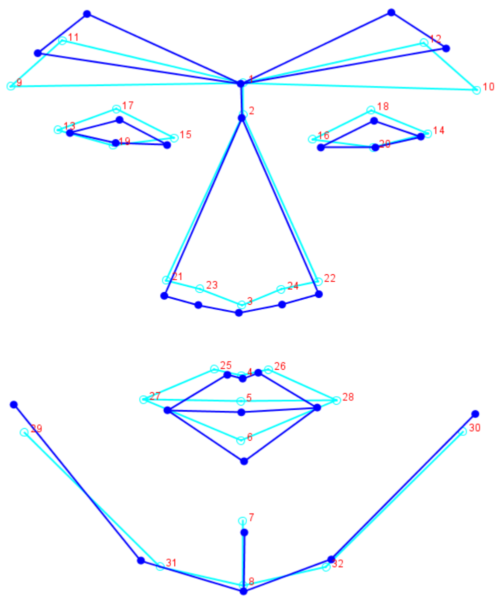


In the wireframe it highlights the elongation of the mandibular corpious can be seen whilst almost all other features remain the same. The mandibular corpius is also a dominant feature that can be used to identify the gender of the individual. The elongation can be used to identify an individual as male or female. With an insight to the dataset and a high number of male individuals I believe the changes we can see are that of males primarily.

Neutral Face - Principal Component 2



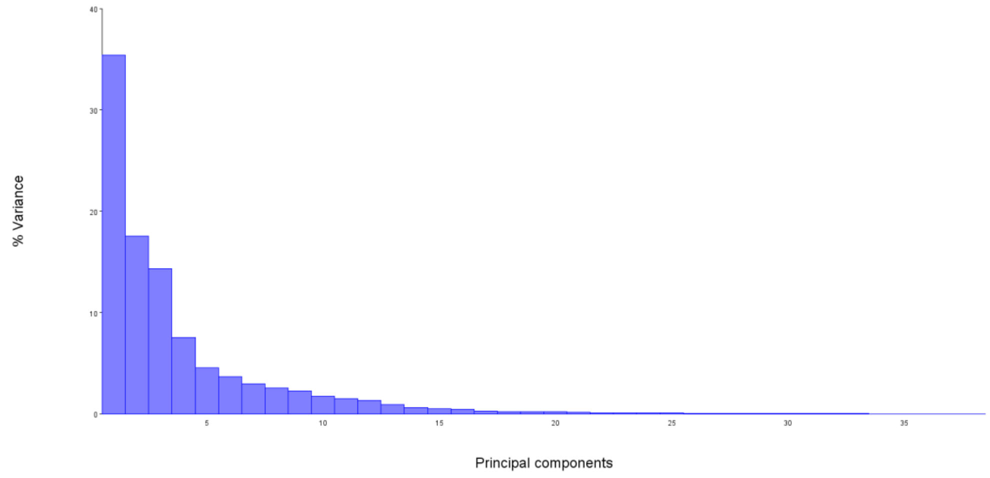
Neutral Face - Principal Component 3



Principal Component 2 shows that all features widen however the jawbone also is a dominant feature that becomes elongated. This principal component shows that the skull becomes larger and all facial features extend with it.

Principal Component 3 show that facial features such as the eyes and mouth change the most. The eyes become thinner whilst the eyebrows become raised and the mouth becomes smaller but elongated. It could be possible that the elongation of the mouth is due to human error whilst land marking.

Eigen values (% variance & cumulative variance)

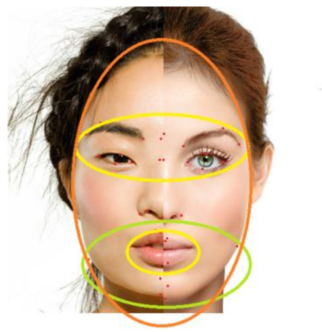


Principal Component Analysis

	Eigen values	% Variance	Cumulative %
1.	0.00269087	35.653	35.653
2.	0.00115756	15.338	50.991
3.	0.00102557	13.589	64.579
4.	0.00047286	6.265	70.845
5.	0.00039332	5.211	76.056
6.	0.00032861	4.354	80.410
7.	0.00024185	3.204	83.615
8.	0.00022410	2.969	86.584
9.	0.00015605	2.068	88.651
10.	0.00014207	1.882	90.534
11.	0.00010854	1.438	91.972
12.	0.00008739	1.158	93.130
13.	0.00007848	1.040	94.170
14.	0.00007029	0.931	95.101
15.	0.00005555	0.736	95.837

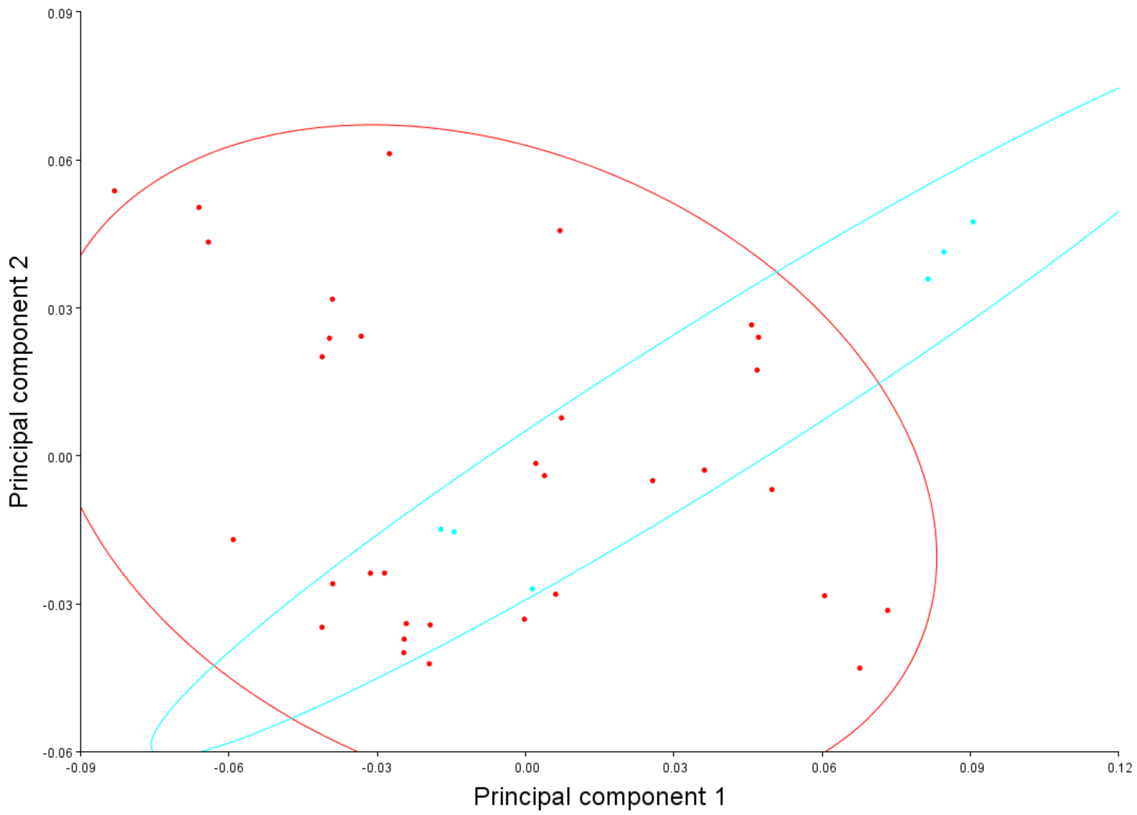
PC1 contributes to approx 36% of the total variance and highlights the mandibular corpious as the most dominant feature presented in the data set. This feature is represented in both males and females and could be an attribute of ethnicity / dietary patterns in a geographical area.

In the example to the right you can see the elongated mandibular corpious and higher brows from principal components 1 **(identified by the green circle in the image).**



PC2 shows the features being extended outwards horizontally and becoming larger **(identified by the orange cicle)**, whilst PC3 shows the brows becoming higher and the mouth becoming larger **(identified by the yellow circles).**

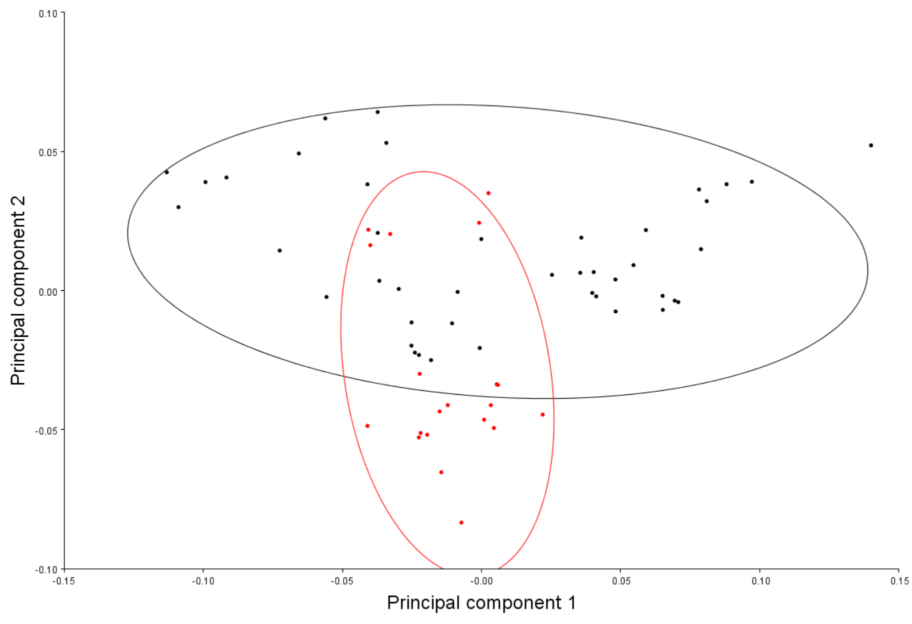
PC Shape Changes - Neutral Faces categorised by gender



The following graph shows that gender cannot be identified with a high degree of confidence however you can predict where females may be on the graph however this area may also represent a male. Even using CVA there is no possible way to represent a clear definition between male and female natural faces.

PC Shape Changes - Neutral Faces categorised by ethnicity

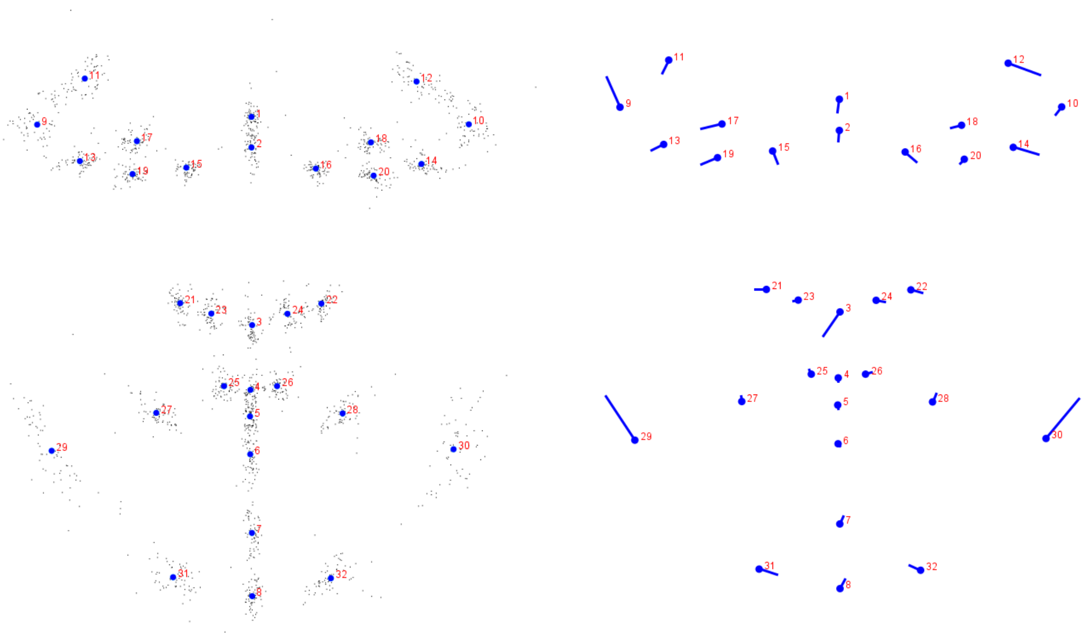
However it becomes much easier to identify neutral faces by ethnicity when we view the confidence ellipses. With the exception of some individuals overlapping it can be seen the Asian ethnicity can be represented by the red ellipse, and Caucasian can be represented by the black ellipse. Viewing ethnicity and the principal components it can be seen that the maximum variance in PC1 is represented amongst Caucasian individuals, whilst PC2 is represented by Asian individuals. This could be attributed to the number of Caucasian to Asian members in the data set.



Analysis of Neutral, Happy and Disgusted Faces with out liars

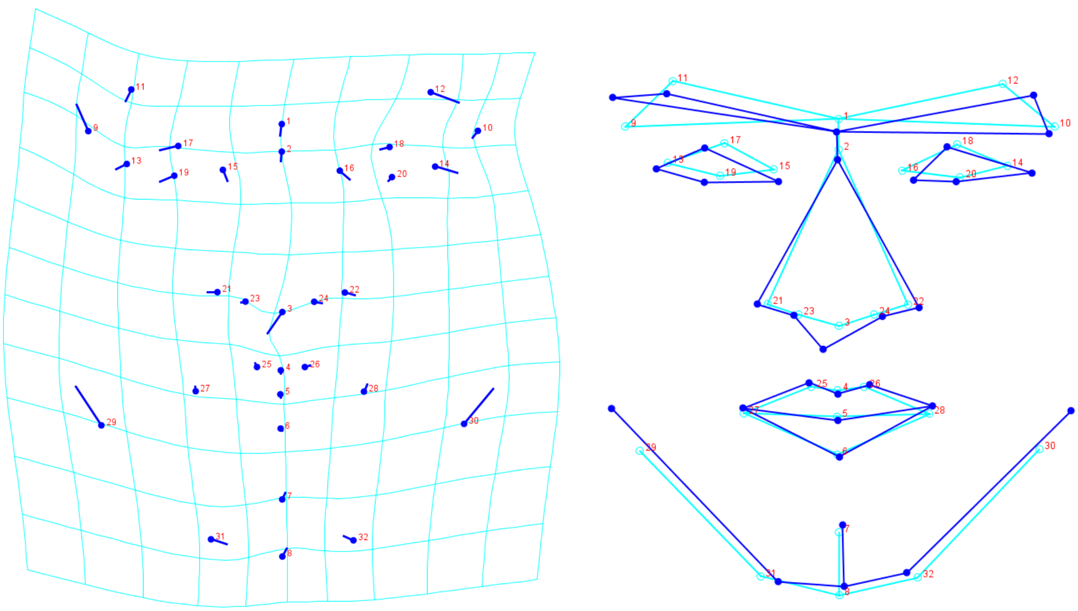
Performing a principal components analysis of the complete data set consisting of three images per classmate: the happy image, the disgust image, and *one of* the neutral images alongside the inclusion of out liars shows how the results can be skewed by a small number of individuals

Lollipop Graph with Outliers



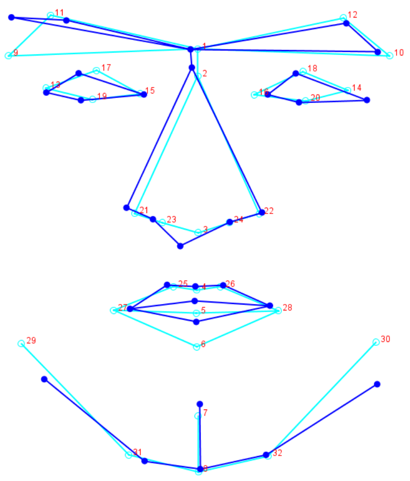
The lollipop graph now shows that dominant features such as the mandible corpious are still represented however now a larger number of features show changes. Compared to the neutral images the eyes are not a key area that change a large amount, however the nose also appears to change which would not be expected due to the limit of movement from the nose.

Transformation Grid and Wireframe with Outliers

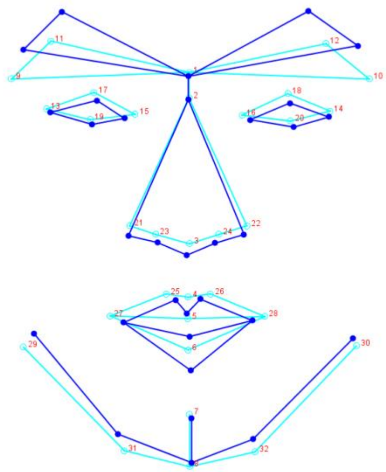


The transformation grid and wireframe now show the extent to which the features can be distorted by out liars in the data set. The most lateral point of the eyebrow has becomes skewed and the most inner point between the nose and the lip has become displaced to the left.

Principal Component 2 with Outliers



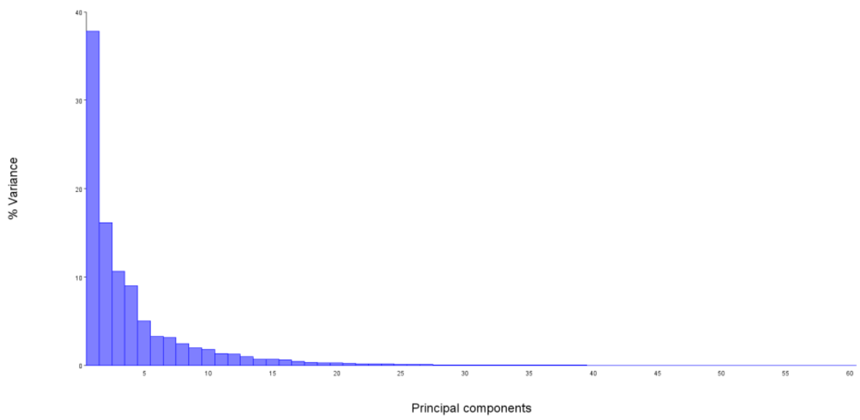
Principal Component 3 with Outliers



Principal component 2 shows the most dominant feature is the change in the most lateral point of the eyebrow and the most lateral point of the right mandible however not these are both on opposite sides.

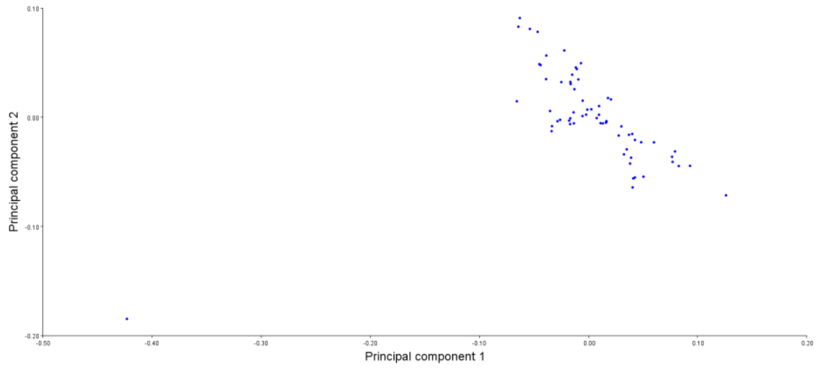
Principal component 3 shows the highest points on the upper margin of the eyebrows become extended upwards but all other features remain the same. It would be expected that the eyes would widen slightly if the brows where to raise.

Eigen values (% variance & cumulative variance)

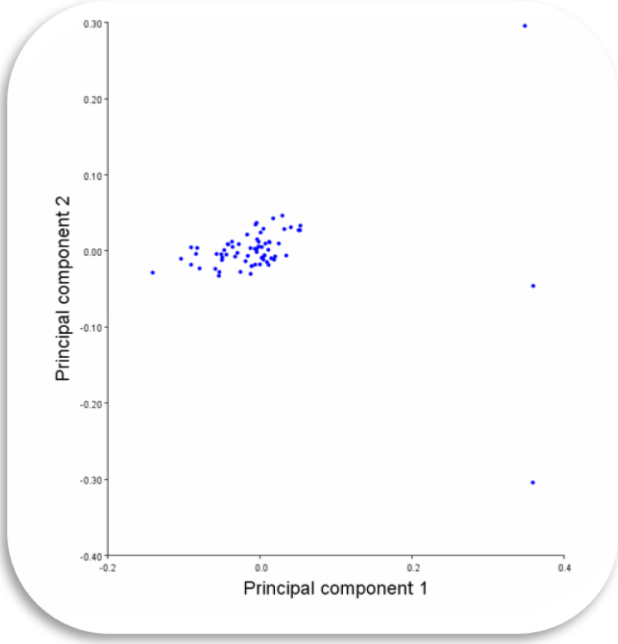


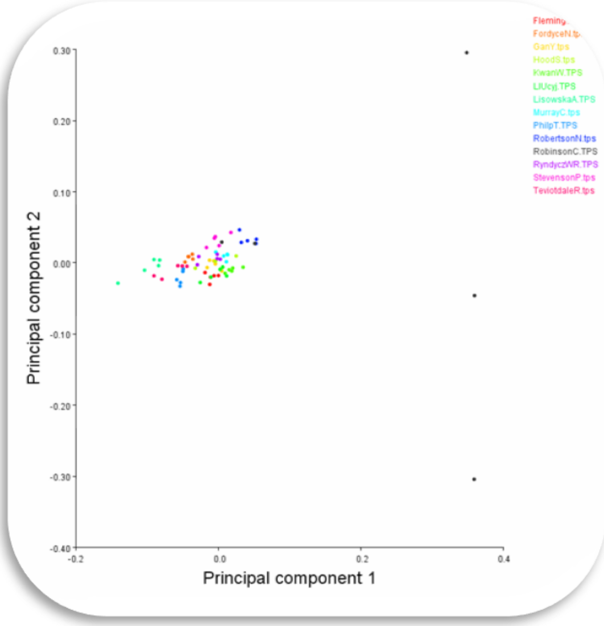
PC Shape Changes

The lower left hand corner indicates that one image is an out liar by a large portion and has greatly affected PC1 and PC2.



Interoperating results when out liars are present

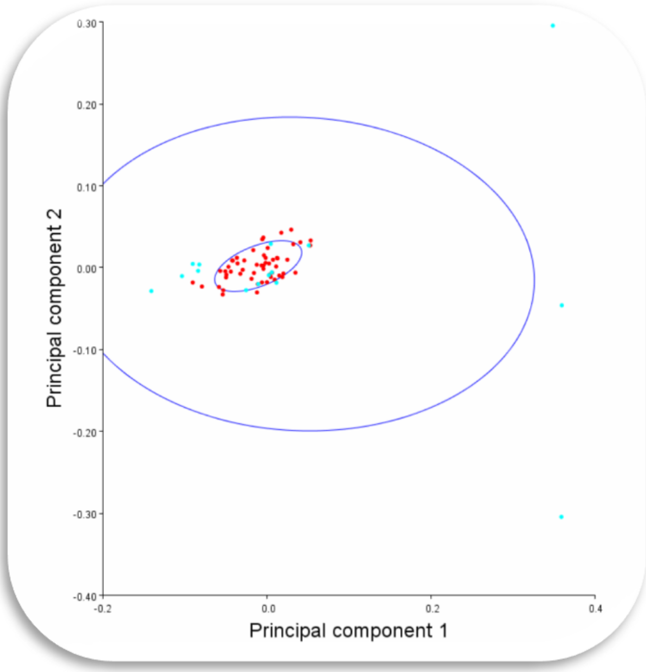




Along the right hand side of the graphs it can be seen that one out liar is very uniform for three neutral faces however these have both greatly affected PC1 and PC2.

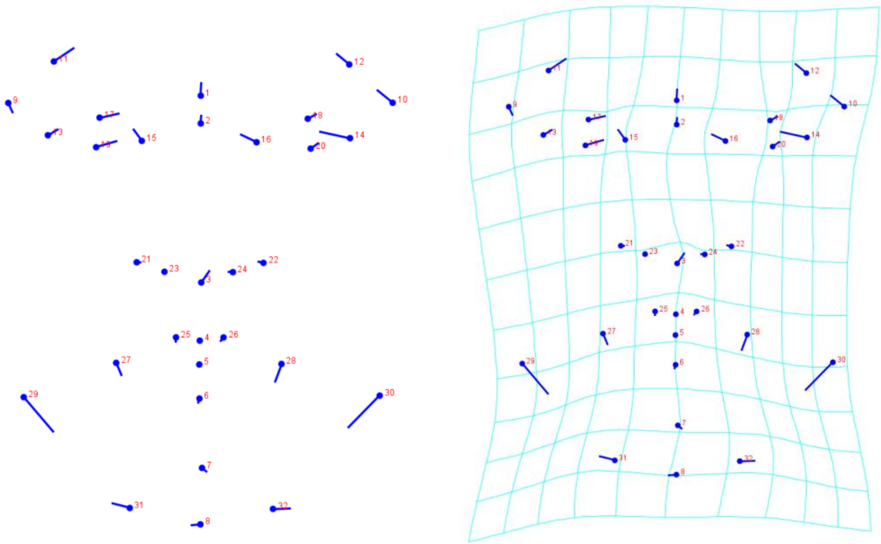
Confidence ellipses for Gender.

The following indicates how out liars can greatly affect the results given. Although females are a small subject of the data set they appear to have the largest variance in both PC1 & PC2. Although this is possible such a small variance between a large number of males would not be expected as some variance would be expected. To interoperate the results at this stage would produce a incorrect final outcome and for this reason the decision has been made to remove all out liars complete from the data set.

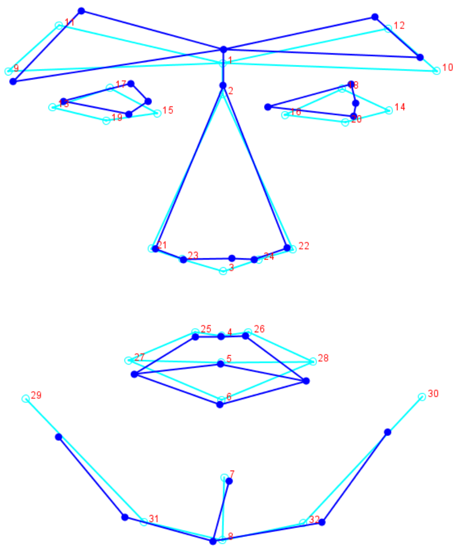


Analysis of Neutral, Happy and Disgusted Faces Without Out Liars

The out liars have been removed from the data set and it now more closely resembles the first data set with neutral faces as the most dominant feature is now the mandible copious once again. The lack of change suggests that the program may still be reading all three neutral faces along with the happy and disgust faces instead of a single one however the settings in the data set appear to have excluded two of them.

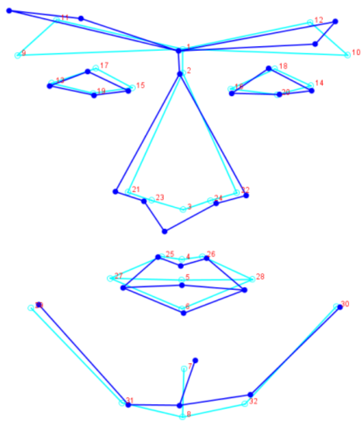


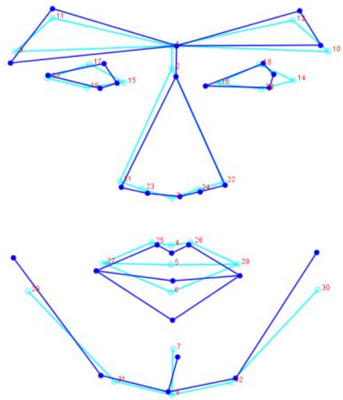
Principal Component 1 Without our liar



The mandible corpious appears to be wider and the brow line also slightly raised and malformed. Some features such as the nose, mouth, and eyes are slightly different indicating there some changes between facial expression however the eyes appear to change shape which may indicate the PC1 has a bias towards the disgust face.

Principal Component 2 without Outliers **Principal Component 3 without Outliers**

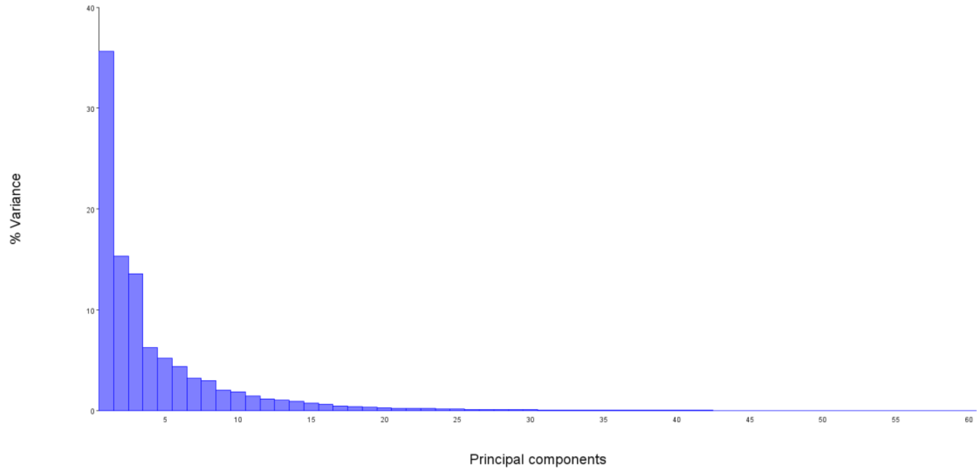




PC2 shows the face widening once again almost as if it was returning to a relax position. I believe PC2 shows a neutral face. Variance in the nose and eyebrows are difficult to account for but this could be a side effect of trying to mimic emotions in a controlled setting and returning to a neutral face afterwards.

PC3 shows the facial features becoming enlarged. I believe this face is the happy face as the eyes appear to widen, and the biggest variance is in the jaw / mouth which would be expected from an individual producing a smile.

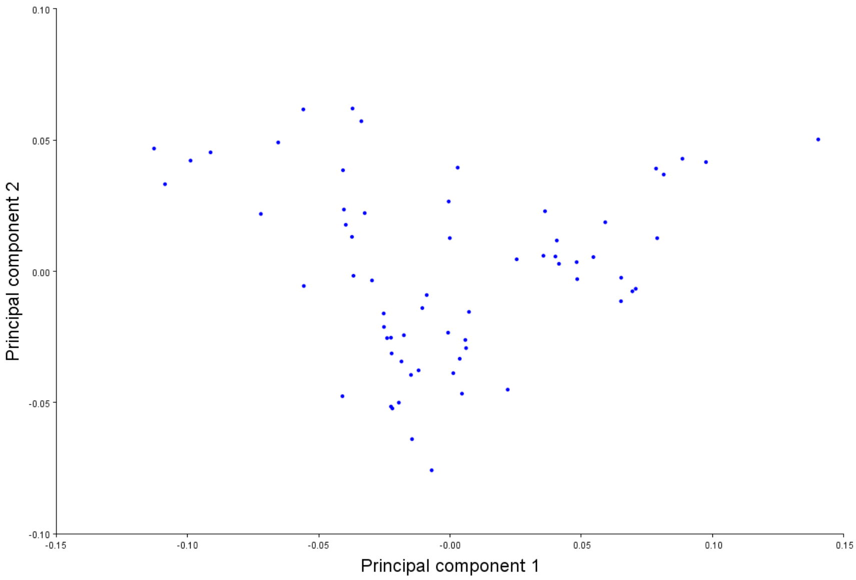
Eigen values (% variance & cumulative variance)



Principal Component Analysis

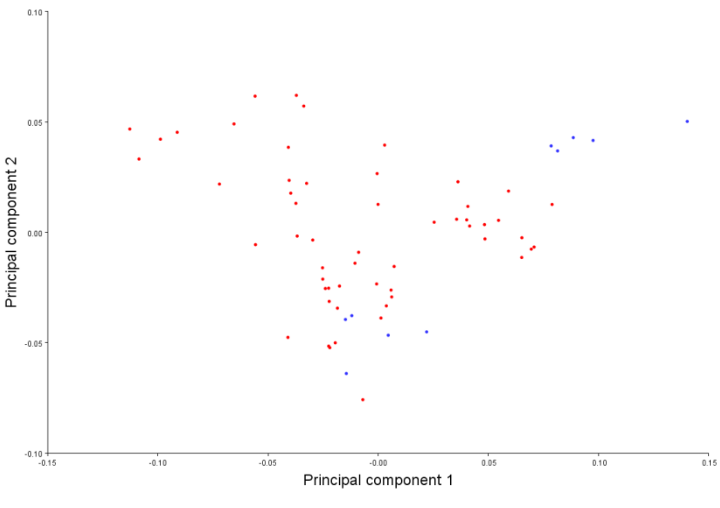
Eigen values	% Variance	Cumulative %
1	0.002691	35.653
2	0.001158	15.338
3	0.001026	13.589
4	0.000473	6.265
5	0.000393	5.211
6	0.000329	4.354
7	0.000242	3.204
8	0.000224	2.969
9	0.000156	2.068
10	0.000142	1.882

PC Shape Changes



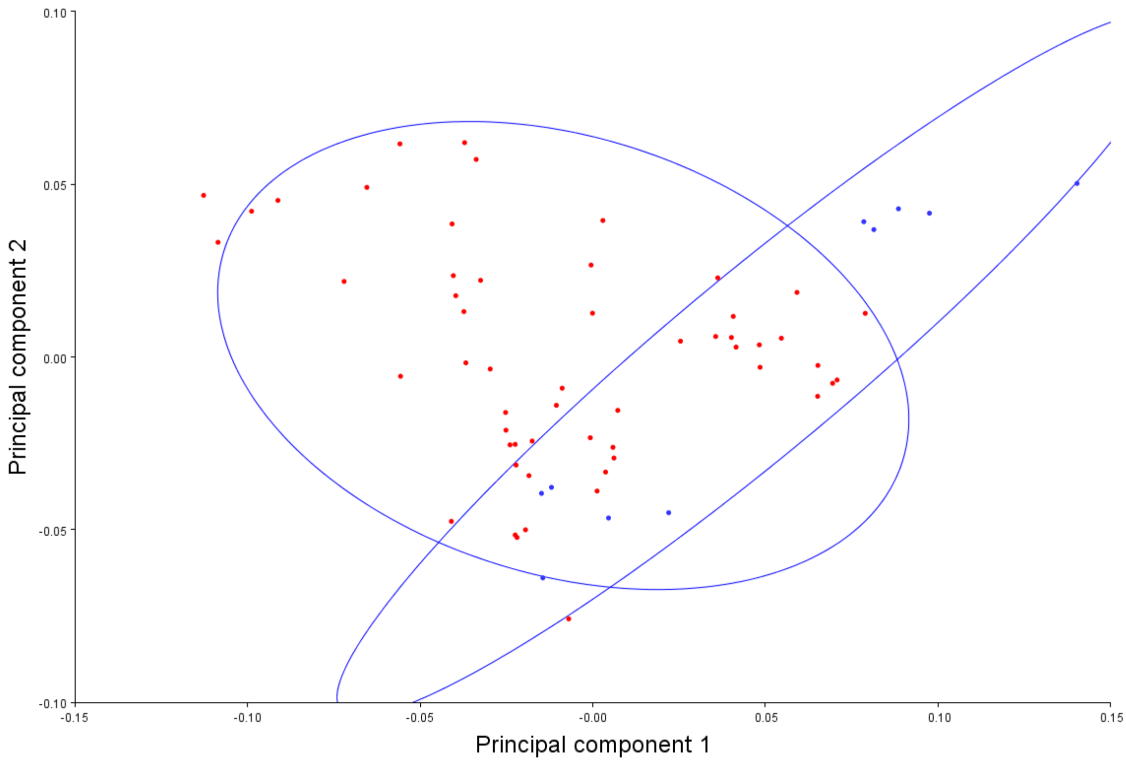
The PC Shape changes can now be seen spread out in almost the same pattern as the neutral faces once the out lairs have been removed. This suggests that to some degree we all produce facial expressions that confirm to the same principal components. Although this may be attributed to a dataset of the same size and range of individuals involved.

PC Shape Changes by Genkder



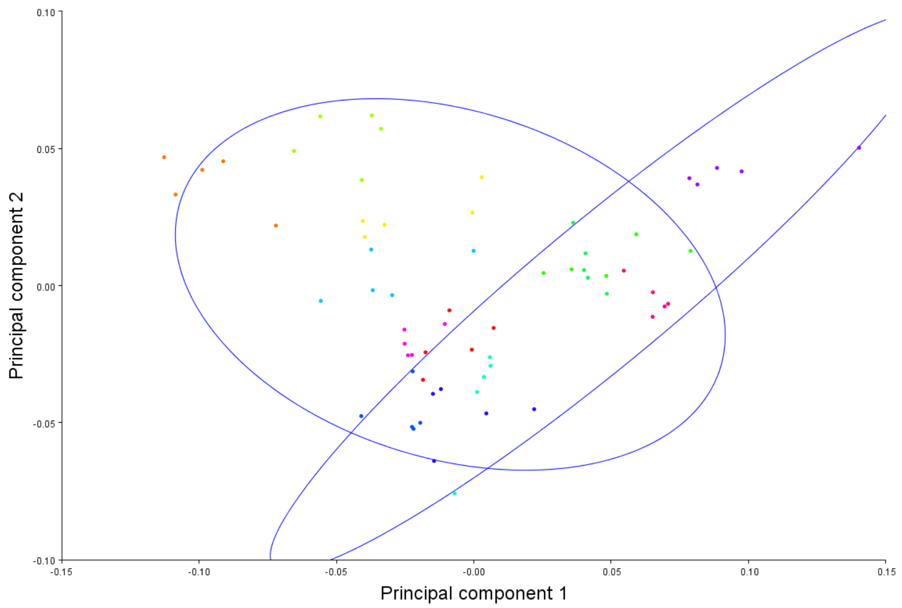
Identifying faces by their gender is difficult due to the small number of females in the data set however using confidence ellipses this can be improved.

Identifying individuals by gender

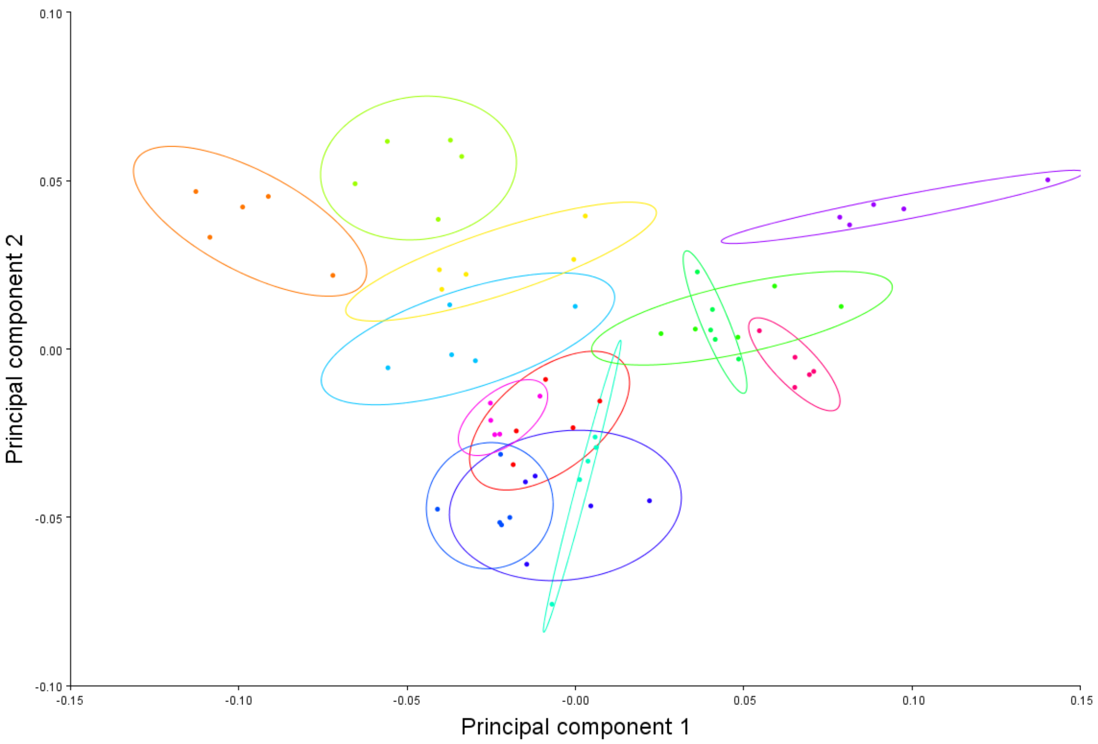


Males are identified by red dots, whilst females are identified by dark blue dots. Although there is some overlap between males and females. It is safe to say that with CVA males and females could be identified with some degree of confidence.

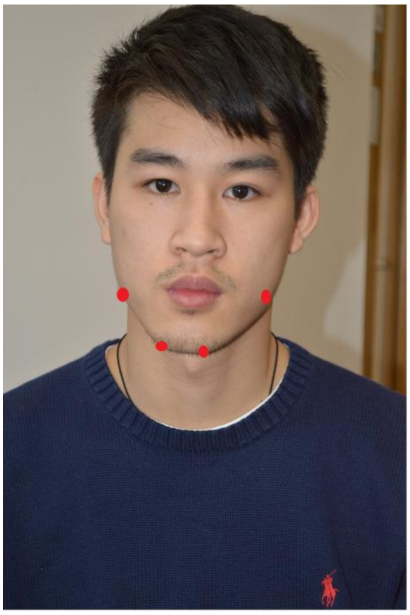
Identifying an individual by their sex



Identifying individuals shows that most individuals facial expressions remain clustered together in a small area. The most variance in PC1 is males whilst PC2 appears to show the greatest variance in females.



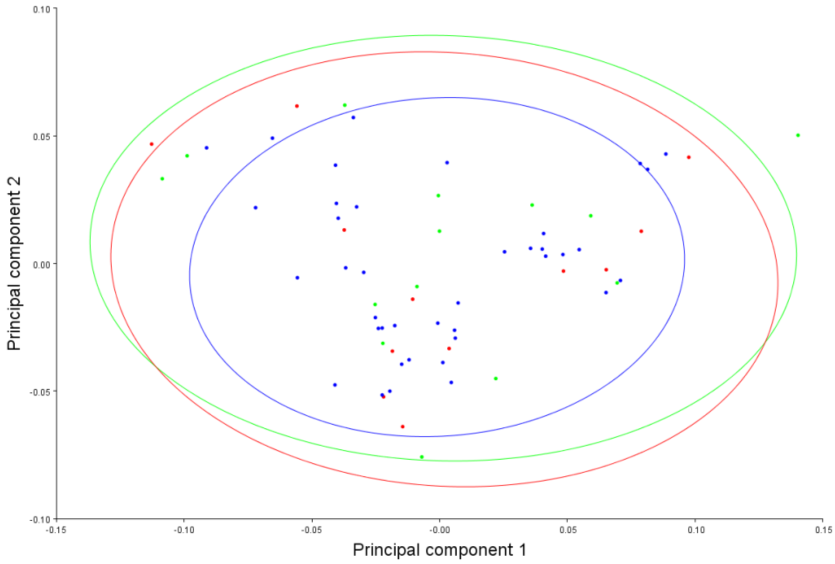
Viewing individuals clearly with confidence ellipses shows that males faces tend to range in variance primarily along PC1 whilst females along PC2. The exception to this rule is Calum Murry, which can be identified by the light blue vertical ellipse. As Calum has a narrower jaw line and PC1 focuses primarily on this then it can be assumed this is why he shows the greatest variance amongst the males. Meanwhile the maximum Variance in PC1 is RnydyczWR.



It is also a possibility that due to a degree of facial hair Calums landmarks may appear closer together effecting his variance in PC2.

Identifying emotions.

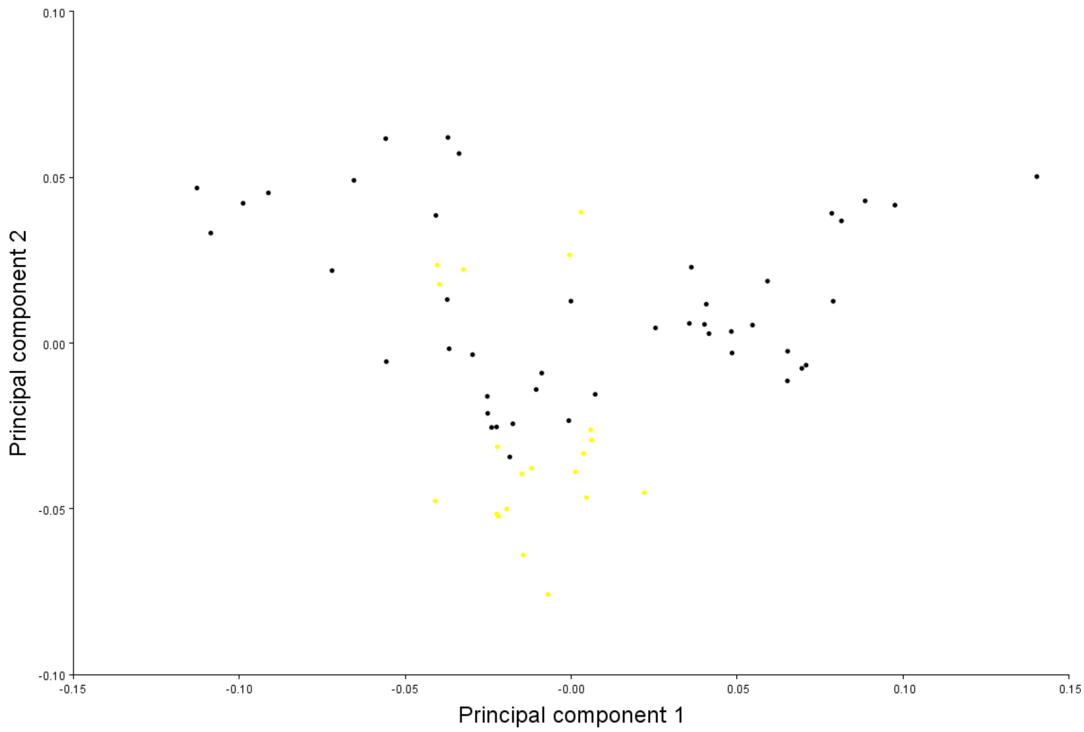
Blue = Neutral, **Red = Disgust**, **Green = Happy**



Attempting to identify emotions cannot be done accurately. The confidence ellipses overlap and have a large variance in both directions. Due to individuals faces clustering together facial expressions would need to be carried out on a per person basis to identify confidence ellipses.

Ethnicity

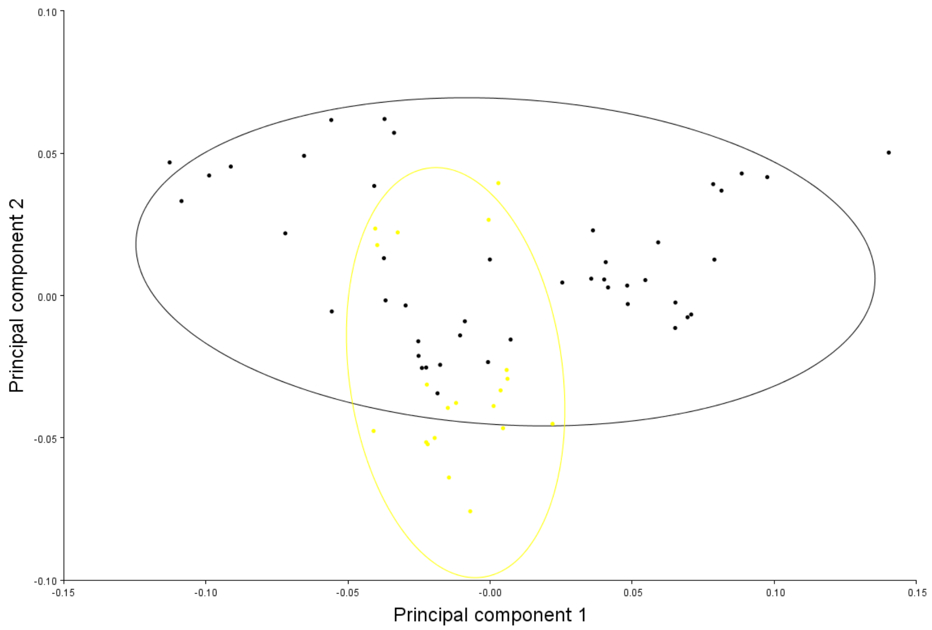
Yellow= Asian, **Black = Caucasian**



Dividing the group into ethnicity such as Asian and Caucasian shows that the largest variance in PC1 is from Caucasians, whilst PC2 is individuals with an Asian ethnicity.

Ethnicity ellipses.

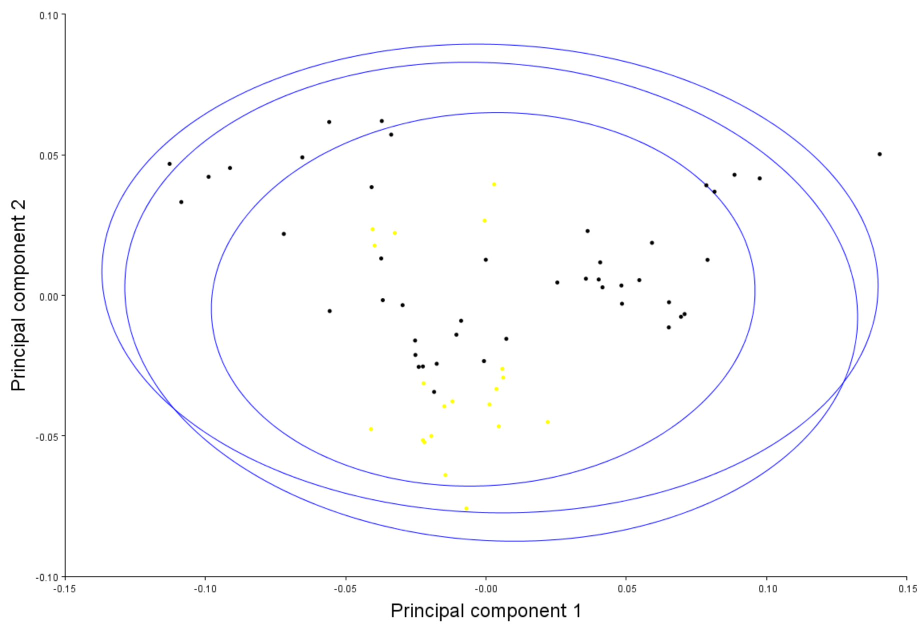
Yellow= Asian, **Black = Caucasian**



With some degree of confidence you can identify where an Asian or Caucasian individual may be. As you proceed along PC2 the likelihood of being Asian increases whilst along PC1 the likelihood of being Caucasian increases.

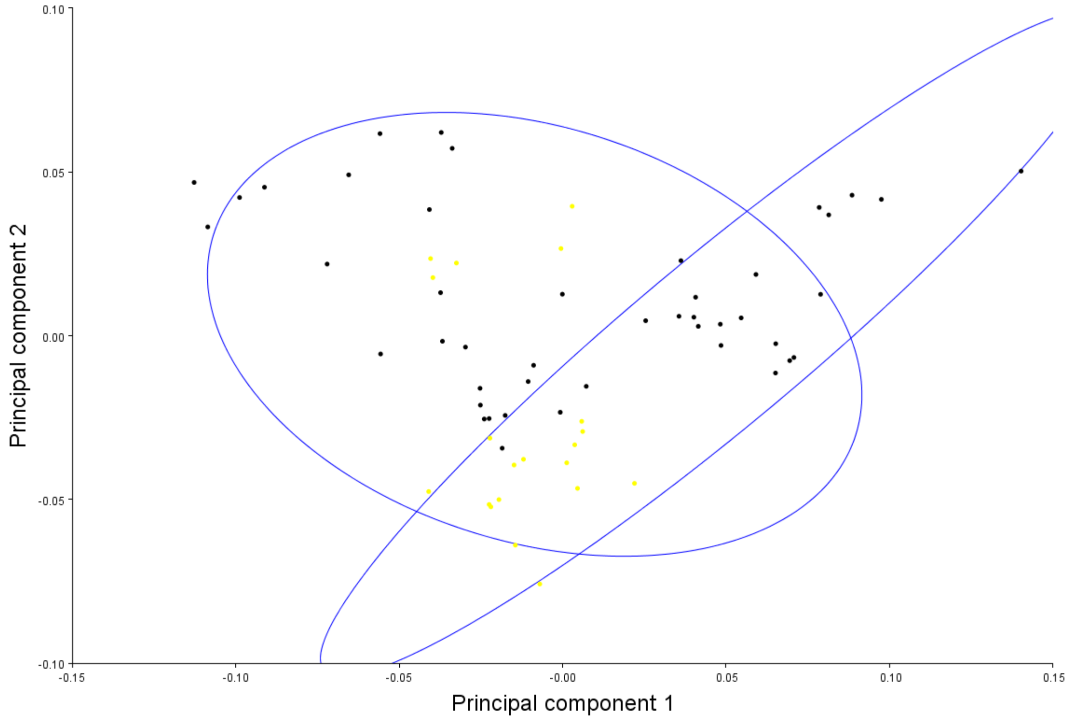
Ethnicity and emotion ellipses.

Yellow= Asian, **Black = Caucasian**



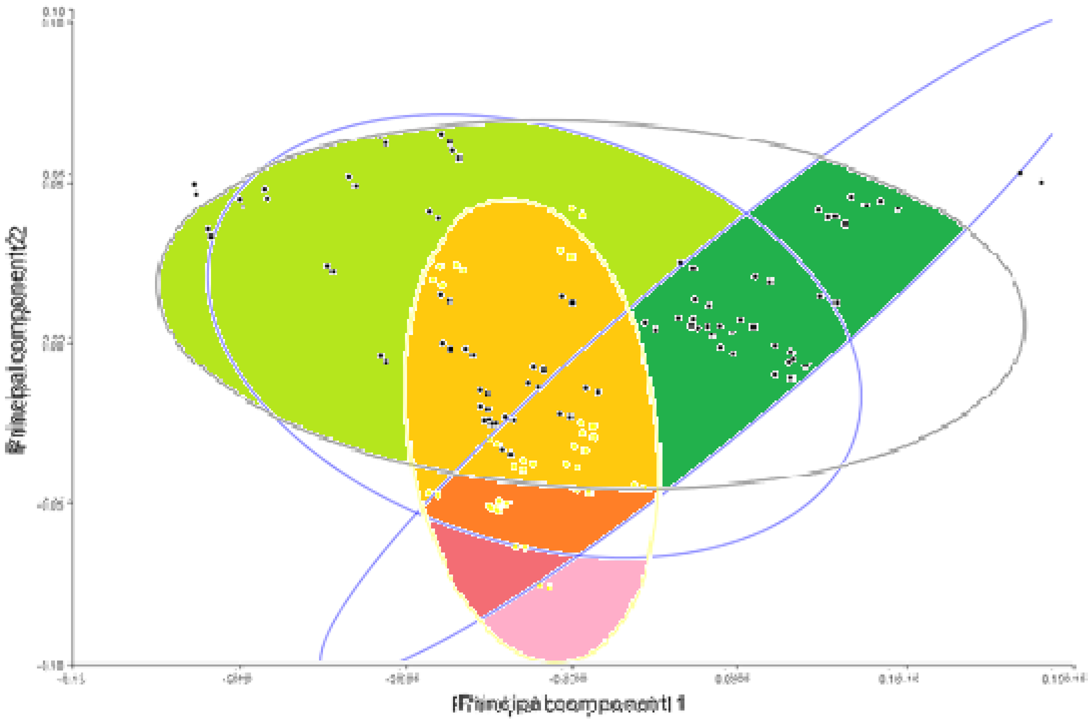
Ethnicity and emotions show no correlation, once again this is due to the clustering of individuals faces to small groups that are distributed across the graph.

Ethnicity by Gender



From this model, without knowing the Ethnicity of a new individual we could determine if they are male / female & white / Asian to a degree of accuracy.

Ethnicity and Gender



x`

The light green section indicates **Caucasian males**, whilst the dark green indicates **Caucasian females**.

The pink section indicates **Asian females**, whilst the dark orange is **Asian males**.

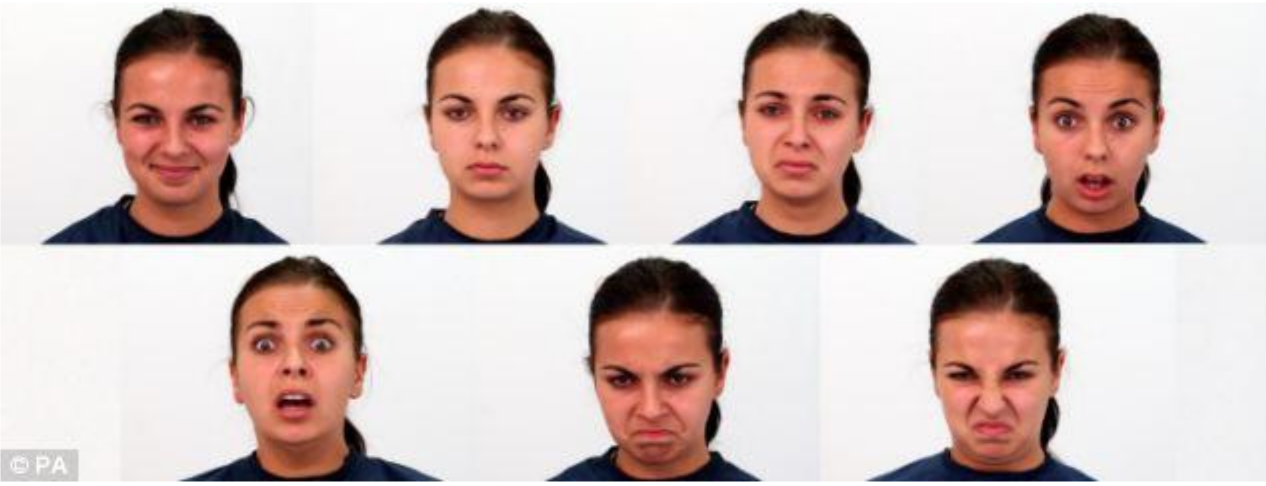
The light orange in the centre is the possibility of **Caucasian/Asian males** and **Caucasian/Asian females**.

Conclusion

The patterns of shape variation in my analysis is that in all experiments, the mandible is the key components to determine the first principal component however this may be simply due to the large number of males in the dataset. The mandible appears to have a correlation with gender / ethnicity to some degree.

In different ethnicities the variance is orthogonal to the other appearing to emphasise other attributes making it difficult to characterise all individuals by a single feature.

During different facial expressions the eyes, mouth, and to some extend the jaw line appear to change and become larger to emphasise emotion. Comparing the neutral faces to a happy / disgust face the happy face becomes enlarged whilst the disgust face becomes smaller. The following image demonstrates this by showing how the facial features of a happy face become enlarged, whilst a neutral face is the basis to compare against. Ally other faces show different emotions however in the disgust face the facial features are scrunched up and matches PC1,PC2, and PC3.



[2]

However to identify these changes clustering of identify has to be taken into account as individuals faces cluster together and a large variance in faces makes these changes difficult to detect in a large dataset. In a smaller dataset it may be possible to determine trends that apply to all individuals.

In viewing certain original landmarked images it is possible to interpretate the results such as in the case of Calum and RnydyczWR showing the maximum variances in both directions. Certain considerations can be made such as if Calums results were subject to noise from the land marking process due to facial hair and even the outline of key features such as the jaw line.

As a whole the dataset makes it difficult to interpreted the facial expressions however on a smaller scale using individuals it may be possible to do this. Clustering of individuals means that their range of facial expressions are subject to the shape of their face and these principal components influence interpreting the results.

.

[2] http://i.dailymail.co.uk/i/pix/2013/03/11/article-0-18902013000005DC-302_634x239.jpg