2D IMAGES CALIBRATION TO FACIAL FEATURES EXTRACTION

Daniel Trevisan Bravo and Sergio Roberto Matiello Pellegrino
Instituto Tecnológico de Aeronáutica, Praça Mal Eduardo Gomes, 50
São Jose dos Campos, São Paulo, Brazil

Keywords: Facial features extraction, digital image processing and computer graphics.

Abstract: The extraction of facial characteristics is important for several applications that involve the 3D reconstruction and facial recognition. In general, facial modelling applications based on 2D images for create 3D model needs, after all, to prepare the extraction images looking for the facial features. In this paper we show some procedures to calibrate and make corrections between two distinct images acquired in distinct instants (the frontal and outline). The aim of this study is to work with human faces, so we use well kwon characteristics like eyes position, mouth, nose curvature, etc. to take previous knowledge about some features and take it to help us to find automatically secondary characteristics for 3D human face. Although this approach makes the process faster, it also imposes restrictions to the model that, however, does not disable its execution in most of cases.

1 INTRODUCTION

The human face has macro characteristics well defined, which become challenging when observed in details. Essentially, the face is the main part of the body and serves as a differential element among the individuals, being responsible for transmitting emotions.

There are many approaches that can be used in the facial features extraction of two or more 2D images that can be used like an automatic or a semiautomatic form (with human interference).

The most part of papers that accomplishes this matter uses two cameras to capture the two images at the same time. It facilitates the process because the images can be previously calibrated. Among them, highlight the works of (Akimoto, Suenaga, and Richard, 1993), (Kurihara and Arai, 1991), (Lee, Kalra and Magnenat-Thalmann, 1997) and (Lee and Magnenat-Thalmann, 2000).

In this paper, the differential is that we take two pictures with the same camera in two different moments. For this, we need to develop a calibration and correction procedures to join the two images to reconstruct the 3D model. This will be discussed on the next sections.

2 2D IMAGES CALIBRATION PROCESS

In this section the calibration process is detailed, what covers, from the restrictions that should be imposed on the acquisition of images, to the final result.

2.1 Images Acquisition

Some restrictions should be adopted during this phase: (i) the image must be in Gray scale, (ii) lightning and environment background homogeneous, (iii) no obstruction of images, (iv) the face must be neutral, (v) the lateral image must be taken, considering that the head must have had a rotation angle of 90° around axis y, (vi) the head inclination angle must not be bigger than 15°and (vii) the images must have a good contrast degree.

However these restrictions were not avoided in all moments, and the results could be obtained in a good form, showing the algorithm's hardness.

Trevisan Bravo D. and Roberto Matiello Pellegrino S. (2007).
2D IMAGES CALIBRATION TO FACIAL FEATURES EXTRACTION.
In *Proceedings of the Second International Conference on Computer Graphics Theory and Applications - GM/R*, pages 124-129
DOI: 10.5220/0002074901240129

2.2 Images Normalization

As the images are captured in different moments, they are not calibrated. The first step in the normalization way of the two images is to find the top of the head and the chin, making this one is in the same vertical position on both images. The next phases are the alignment of the others features like mouth, nose and eyes.

To normalize the images we need to align the constellations of the frontal images with the lateral one.

To help the speed of the process, we used a facial algorithm characteristic location, developed by Coelho (Coelho, 1999). It helps with the reduction of the searching area, finding the rectangular region that recovers the head in the front view. The same algorithm was used for the lateral view, but it had to be adapted for this case.

We look for the tip of the nose at the first step, because it is the most salient point in the lateral view, and it helps us to define the right limit (for example) of our model. The opposite limit is not easily found, because this one can be camouflaged by voluminous hair; the tied hair extends the head length, and the others. In this case, it is defined as being the horizontal smaller position found in the central head height (the mid point from the head until the chin). The inferior limiting is found by the profile outline topography analysis (Bravo, 2006).

The analysis is made by extracting the outline lateral profile limit, as shown in the Figure 1 (a). As the image background is homogeneous, the edge that represents the side profile is detected through of horizontal and vertical scan lines running through the image from the top of the head to the neck region. The point is considered belonging to the edge when there is a pixel luminous intensity abrupt variation, thus, the profile outline is obtained, besides being used to the determination of the inferior limiting and in the features extraction phase.

To determine the inferior limiting, a profile topographical analysis is accomplished in a region from the nose height until the neck. In Figure 1 (b) it is shown the topography analysis in a part of outline, where *α1*, ..., *αn* are the angular coefficients that the profile does outline straight line with the axis *x*.

We can see that the most negative coefficient always coincides with the inferior limit of the face, so this information leaves us to conclude that this point corresponds to the smallest Y value of the image.

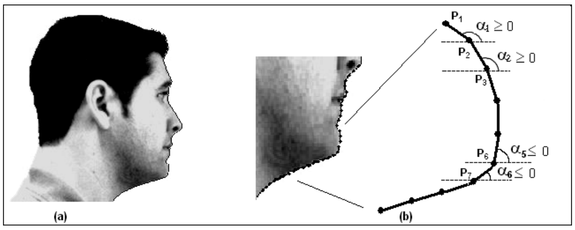

Figure 1: (a) Edge that represents the profile outline; (b) Profile outline topography analysis used in the detection process of the inferior limiting.

To calibrate the images we insert them in two boxes with pre defined size, keeping the height and width proportion between each other. The bounding boxes are showed in the Figure 2.

Figure 2: Determination of the rectangular regions that recover the head.

The algorithm looks for the objects of interest using the techniques of Constellations, which consists in finding the group of pixels with low intensity. In this way, five Constellations are found, the eyes, the nostrils and the mouth in the front image, and three in the lateral one. In the last case, we adopt an empirical solution to find the new area where the Constellations are. First of all, we delimited the area of search using the following criterions: The superior limit of the new box is defined as the mid point between the top of the head and the tip of the nose; the inferior limit is taken as the own chin position; The right border is taken passing on the tip of the nose; finally the left border are far away from 80% of the left side of the bounding box.

Each constellation is composed by the minimum points (marked in white) and by the central point (marked with a cross), as it shows the Figure 3 (a).

It can have some constellations that do not bring interesting information. When you know the tip of the nose location, it can say that, above of this point, there will only be the constellation of an eye and, below it, only two constellations, of the nostril and of the mouth. With these information, the lateral image constellations are defined (Figure 3 (b)).

After all Constellations discovered in the two images, the system proceed the verification if they have the same vertical position, using the information of the central point of them like target. If they are not on the same height, we can tolerate some bypass if the central point of one is inside the range of the other Constellation.

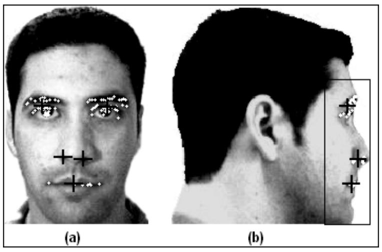

Figure 3: (a) Constellations and minimum points defined for the frontal image; (b) Constellations and minimum points defined for the lateral image.

When all the constellations are aligned, the normalization process will be completed. In Figure 4 there are shown the normalized images.

Figure 4: Normalization process of the images through the constellations aligning.

It can happen that, one or more constellations are not aligned. In this case the problems can come from the images acquisition phase. There is a possibility of a no alignment of all constellations due the head inclination to be different at each acquisition moment. When it happens, the generic face model is applied to the images, in a way that, for any alteration on the front mask, an alteration in side image is provoked, inserting or removing lines in the original one.

3 FEATURES EXTRACTION

The features are extracted in both images following the principle of “divide and conquer”. The information is obtained performing the image analyses in each region separately instead of scanning the whole face. It accelerates the work, mainly in relation to the processing time to get to a final result. In Figure 5 the points to be extracted are showed, and are based in a subset of feature points determined by the MPEG-4 pattern. More information about this pattern can be obtained in (Pandzic and Forchheimer, 2002).

Figure 5: Feature points to be extracted in frontal and side images.

In Figure 6 one can see the images with the defined search windows. The dimension of each one is calculated based on the central point of constellations, being the length and height defined by percentage in relation to the face dimensions on both views.

The search procedures by feature points in those windows employ some specific techniques of processing and analyzes of the images.

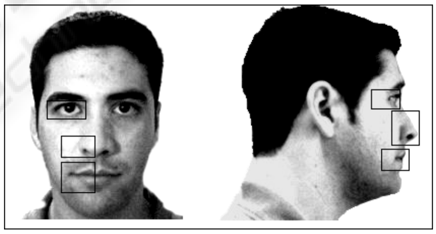

Figure 6: Defined search windows to features extraction.

3.1 Feature Extraction on the Front Image

Considering that the human face is almost symmetric, the features extraction on the front image is applied only for the left half of the face.

The features extraction of the eye region is based on the work developed by (Vezhnevets and Degtiareva, 2004), who uses procedures based on the luminosity intensity variation of the pixels.

In Figure 7 the extract process of the eyes extremities, are divided in four steps: 1) to detect the centre of the iris; 2) estimative of the iris ratio; 3) to detect the border of the inferior eyelid; 4) the estimative of the superior eyelid contour. The determination of the points that belong to the border of the eyelids, enables the features extraction that gives information about the height and width of the eyes.

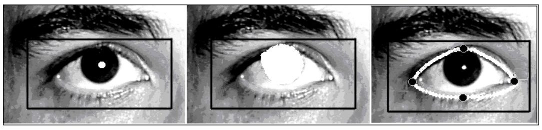

Figure 7: Eye features extraction process.

Having the line segment which unites the eyes extremities and the line segment defined by the inferior and superior limits, and which passes through the tip of the nose, the central axis of the face must have an inclination of 90° in relation to axis x (Figure 8).

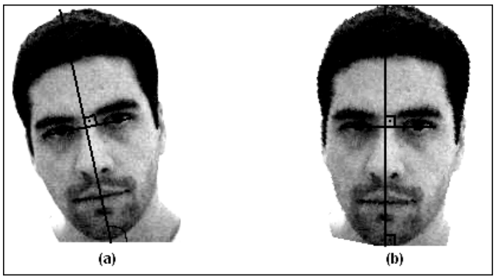

Figure 8: (a) Central axis determination; (b) Head inclination correction.

In Figure 9 the other features extracted for the nose region are showed. The coordinate *y* of those points are obtained having as reference the face model of Parke (Parke and Waters, 1996). The coordinate *x* of the point $N1_F$, represents the nose edge, and are determined applying a Sobel Operator (Lindley, 1991), in the same way the coordinates of $N2_F$ helps us to find some peculiarity of the nose form like its root on the front, the curve near the corner of the eyes, and the silhouette from the middle up to the base in the down side of it

The application needs extracts the features from the mouth region. For that we are based on the work developed by (Lanzarotti et al., 2001). The first step is detected the cut line between the superior and inferior lips. When the system knows this region its enable to extract the points related to the mid point of the lips and to the corner of the mouth respectively, $M1_F$ and $M2_F$, as showed in Figure 10(a).

Moreover, it is possible to classify the mouth region in three parts (Figure 10(b)): Superior lips (A); Inferior lips (B); Central of the mouth (C). This way helps the system to identify more easily the own characteristics, because it will work in specific area.

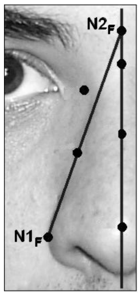

Figure 9: Features extracted for the nose region.

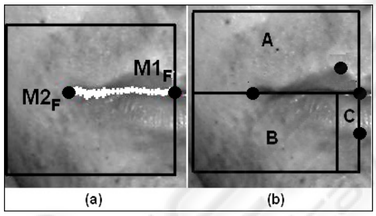

Figure 10: (a) Determination of the points M1F and M2F by the border of the lips; (b) Subdivision of the mouth region to the extraction of the related points to the superior and inferior lips.

In order, to extract the target points, the face are divided in two regions as showed in the Figure 11(a). In the same figure, part b, we can see the target points seek in the region I, correspondent to the front of the model. Their locations are based in the Parker model, and it's taken as help. In the regions II the points that will be used in chin silhouette are founded. The extraction process of them is based on (Goto et al., 2001), where the jaw curve is determined taking into account that its shape is approximately half of an ellipse. The Figure 11(c) shows the points which are extracted from the outline of the jaw and the information of borders on this region. Such border is obtained by a searching system in a region limited by two auxiliary ellipses, one internal and another external to the ideal curve.

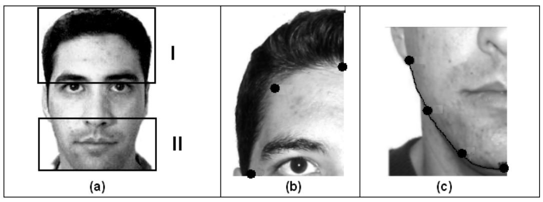

Figure 11: (a) Face division for feature extractions; (b) Extraction of the points related to the front head region; (c) Extraction process of points on the jaw border.

Some points belonging to the chin region will be obtained when its corresponding ones are extracted from the side image.

3.2 Feature Extraction on the Side Image

The points extracted in this phase concerning the vertical and horizontal position (depth), respectively *y* and *z* coordinates.

For the eyes it's only possible to extract three points: the left corner and the superior and inferior limits, as showed in the Figure 12.

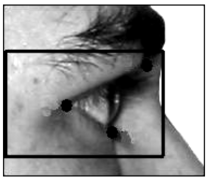

Figure 12: Features extracted in the eye region.

In the nose region we have some interested points. The *N1s*, *N2s* and *N3s* points are easily extract with help of the other found earlier. The other points showed in the Figure 13 are obtained creating two sub regions, where the border passes exactly in the tip of nose.

From sub-region B are extracted the points referent to the nose edge ($N4_S$) and the nose base ($N5_S$). These data are obtained with the help of extractions done on the front face, being the relation shown in Figure 13 (b).

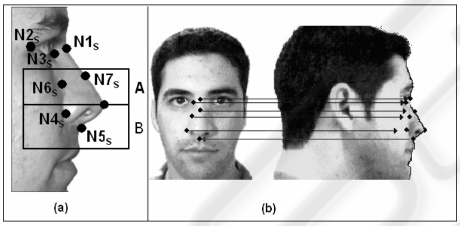

Figure 13: (a) Subdivision of the search window for extraction; (b) Correspondence between the images for the feature extraction in the region of the nose.

In the sub-region A, the vertical positions that characterize the curve that narrows the nose ($N6_S$ and $N7_S$) are obtained through the features correspondent on the front view. The coordinate *x* of the point $N7_S$ is obtained considering the border of the outline at the height where the feature was defined.

The process of extraction in the mouth region is based on the work developed by (Ansari and Abdel-Mottaleb, 2003). First of all, we sub divide the region in two parts with purpose to identify the two more external points of the silhouette. The border of this two windows passes through the point *M3s*, showed in the Figure 14(a), and it was obtained from the cut of lips

After that, we can easily obtain the points *M1s*, in the window *A*, and *M2s*, in the window *B*. The last interested point (*M4s*) is the corner of mouth, and is obtained from the information of the borders in the region limited by search window, as showed in the Figure 14(b).

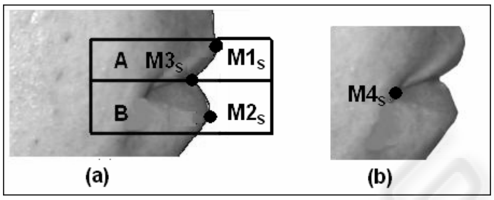

Figure 14: (a) Subdivision of the search window for extraction; (b) Feature extraction of the mouth border.

Finally the system will detect the contour of the face in the lateral side. For this case the image was divided in two parts, showed in the Figure 15. To obtain the inner contour we use an ellipse to help us to find the interested points. The outer points in the chin are extracted normally from the face border information.

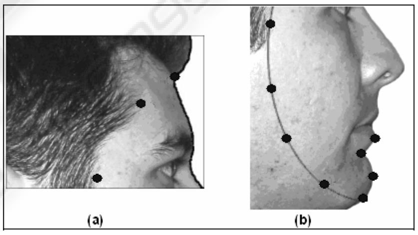

Figure 15: (a) Extraction of the forehead features; (b) Extraction of the features of the face silhouette.

4 RESULTS

We analyzed 72 images of 33 people. The images database generated during the research looked for to obtain the biggest possible variety of physical types without worrying about ethnics.

In normalization phase, 90% of the analyzed images could be normalized. Some images could not be calibrated because the algorithm used did not detect the constellations properly in case of beard or when an image had a little contrast. The accuracy in detecting the limits of the windows that recover the images also presented fault.

The extraction process of the feature was showed efficient in approximately 80% of the cases. In a small quantity of images, the features were not correctly extracted, because some characteristics as fringe on the forehead, glasses and the head inclination disarray the system.

Based on many tests, anyone could see that each face region geometry vary a lot, mainly on the forehead, chin and nose regions. In these situations, some changes were made on the adaptation model. We saw also, that the use of glasses can distort the geometry in the nose region, once the features are not extracted adequately on the side image.

5 CONCLUSION

This study showed the real possibility to acquire two images in different moments and process their to obtain a realistic 3D model of a human face. We also could see that the illumination homogeneity is important, but not fundamental. We have got good results without any additional care about it.

The real possibility to do all process automatically could be proved. The only human interference occurs in the initial moment to adjust the window to limit the model head, all the other process are made automatically and rapidly.

The time consuming for all execution procedure is not greater than 32 seconds. This efficiency comes from that the system knows in advance which characteristics will be searched, decreasing the computational effort.

In order to make the process happen without human interference, it was developed a correction algorithm for the head inclination, which provided and excellent alignment between the two images, making the work of feature research. We should highlight that the algorithm has its limitations and, for inclinations over 30%, it can fail, however in normal conditions, it was showed efficient and it did not accumulate errors to the synthetic images.

Moreover to extract the features correctly, it is advisable that the models not present any element that obstructs some regions of the face.

An improvement of this work would be the identification of elements, such as, the use of glasses, fringe and beard. With this recognition, such occurrences could be eliminated and make the process possible.

REFERENCES

Akimoto, T., Suenaga, Y., Richard, S.W., 1993. Automatic creation of 3D facial models, *IEEE Computer Graphics and Applications.*

Kurihara, T., Arai, K.,1991. A transformation method for modeling and animation of the human face from photographs. In *Proceedings of Computer Animation'91.*

Lee, W., Kalra, P., Magnenat-Thalmann, N., 1997. Model based face reconstruction for animation. In *Proceedings of MMM'97.* World Scientific Press.

Lee, W., Magnenat-Thalmann, N., 2000. Fast head modeling for animation. *J. Image Vision Computer.* 18 (4).

Coelho, M.M., 1999. Localização e reconhecimento de características faciais para a codificação de vídeo baseada em modelos. *Master's dissertation, ITA, São José dos Campos.*

Bravo, D.T., 2006. Modelagem 3D de faces humanas baseada em informações extraídas de imagens bidimensionais não calibradas. *Master's dissertation, ITA, São José dos Campos.*

Pandzic, I.S., Forchheimer, R., 2002. *MPEG-4 Facial Animation: the standard, implementation and applications,* John Wiley & Sons. England.

Vezhnevets, V. et al., 2004. Automatic extraction of frontal facial features. In *6th Asian Conference on Computer Vision.*

Parke, F.I., Waters, K., 1996. *Computer Facial Animation,* AK Peters. EUA.

Lindley, C., 1991. *Practical image processing in C: acquisition, manipulation, storage,* Jhon Wiley & Sons. New York.

Lanzarotti, R. et al., 2001. Automatic features detection for overlapping face images on their 3D range models. In *11th International Conference on Image Analysis and Processing.*

Goto, T. et al., 2001. Facial feature extraction for quick 3D face modeling. In *Signal Processing: Image Communication*.

Ansari, A-N., Abdel-Mottaleb, M., 2003. 3-D face modeling using two views and a generic face model with application to 3-D face recognition. In *IEEE Conference on Advanced Video and Signal Based Surveillance.*