Generating facial emotions using anthropometry and formal languages for individuals suffering from psychiatric disorders

Rafael L. Testa*, Antônio H. N. Muniz*, Carpio L. U. S.*, Rodrigo S. Dias †, Cristiana C. A. Rocca †,
Ariane Machado-Lima* and Fatima L. S. Nunes*
**School of Arts, Sciences and Humanities*
University of São Paulo, São Paulo - Brazil
Email: rafael.testa@usp.br, antonio.muniz@usp.br, liseth.segundo@usp.br, ariane.machado@usp.br and fatima.nunes@usp.br
†*Department of Psychiatry - Medical School*
University of São Paulo, São Paulo - Brazil
Email: rdgdias@gmail.com and crisrocca@gmail.com

***Abstract*—The ability to process and identify facial emotions is an essential factor for an individuals social interaction. There are certain psychiatric disorders that can limit an individuals ability to recognize emotions in facial expressions. This problem could be confronted by making use of computational techniques in order to develop learning environments for the diagnosis, evaluation and training in identifying facial emotions. This paper presents an approach that uses image processing techniques, formal languages, anthropometry and Facial Action Coding System (FACS) to generate caricatures that represent facial movements related to neutral, satisfaction, sadness, anger, disgust, fear and surprise emotions. The rules that define the emotions were determined using an AND-OR graph to enable generating these images in a flexible manner. An evaluation conducted with healthy volunteer users showed that some emotions are more easily recognized, while for other emotions the caricatures need to be further improved. This is a promising approach, since the parameters used provide flexibility to define the emotional intensity that must be represented.**

***Keywords*-AND-OR Graphs; facial recognition; facial emotions; facial expression; anthropometry; computational grammars; splines; formal languages;**

I. Introduction

The capacity to process and identify facial emotions is an essential factor of human communication and social interaction according to cultural norms and customs, based on common standards. Six universal emotions have been mapped, namely; satisfaction, sadness, anger, disgust, fear and surprise, which correspond to a specific pattern of facial muscles, with partially separable neural circuits that can prepare an appropriate behavioral response [1], [2], [3].

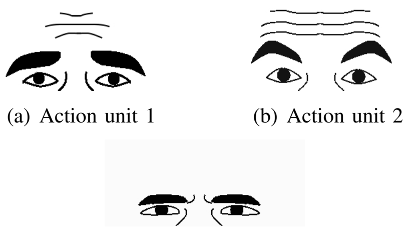
With regards to psychiatric disorders, the face recognition paradigm has received more attention, especially in Autism Spectrum Disorders (ASD) [4], mood disorders [2] and schizophrenia [5]. Computational techniques that include image processing, computer graphics, pattern recognition and virtual reality have been closely associated with developing tools for the diagnosis, evaluation and training in the recognition of emotional facial expressions [6], [7], [8]. What has been verified, however, is that most of these tools are limited in terms of flexibility to generate an adequate variety of objects. Such flexibility is necessary when the goal is to provide autonomy to healthcare professionals so they can adapt the tools for diagnosis or for training in a specific situation or for a particular patient.

This paper describes the initial results of a project intended for developing expressions to represent different intensities of facial emotions assigned for the training and rehabilitation of individuals with mental disorders. The grammatical techniques (AND-OR graphs) used detailed the initial representation of the images that represent expressions of the six emotions cited before.

The caricatures were constructed based on the concepts of anthropometry and Facial Action Coding System (FACS), presented with the AND-OR graphs definition in section **??**. Related works are presented in section II. The methodology used is presented in section III. Section IV describes the results obtained with the approach presented, exposes the discussion achieved and the conclusions about the proposed approach.

Anthropometry is the science that studies the measurements of the size, weight and proportions of the human body, from the measurements of a population [9]. These measurements were determined by a set of well-defined points on the human body, called *landmarks* points. The distance be-tween two landmarks characterizes an anthropometric measurement. Based on the universality of facial emotions, this paper determined techniques to relate the measures across parts of the face with the facial movements involved in the expressions generated by emotions. In [10] the authors present the facial expressions of emotions considering the fundamental actions of a muscle or group of muscles, called Action Units (AUs). Thus, a facial expression is represented as a combination of these Action Units. [11]

In order to illustrate the Actions Units, Figure 1 shows the AUs 1, 2 and 4. In action unit 1 (Figure 1(a)) the inner portion of the brows is raised; in AU 2 (Figure 1(b)) the outer portion of the eyebrows is raised and in AU 4 (Figure 1(c)) the eyebrows are lowered and drawn closer.



(a) Action unit 1 (b) Action unit 2

(c) Action unit 4

Figure 1. Example of Action Units 1, 2 and 4.

Spline is a mathematical curve defined by two or more control points. The control points create smooth curves and then shape its form. Thus, the definition, modification and manipulation of spline curves are carried out through changes in its control points [12].

An AND-OR graph is a sixfold representing the grammar of the image $G =< S, V_T, V_N, R, C, P >$, where S is the initial symbol of the grammar, V_T represents the terminal nodes and V_N represents the set of non-terminal nodes, R is the set of production rules productions, C is the set of all valid configurations that may be derived from the G, and P represents the probability model defined in G [13]. Graphically, this grammar is described by an AND-OR graph in which the root node represents the initial symbol and the leaf nodes represent the terminal symbols, the internal nodes represent non-terminal symbols and the vertical edges are related to the productions: the children of an OR node indicate all possible productions for the same non-terminal symbol, and the children of an AND node indicate all symbols that are generated by a particular production [14].

II. Correlated works

The studies aimed at generating virtual images expressing facial emotions make use of game engines or virtual reality ready packages [15], [16], [17], [7], [6]. However, with regards to flexibility the majority of these resources are limited in generating an adequate variety of emotional intensities and characteristics of avatars such as gender, ethnicity and age, including the limited effortlessness in changing such features to implement these parameters in the game.

A system for creating three-dimensional anthropometric models of the human face is described in [18]. In this work, a set of anthropometric measures are generated to compose different basic models and then these various basic models are employed to generate several combinations of possible face representations.

The article [19] uses AND-OR graphs to represent the aging of faces. The proposed AND-OR graph has three levels: the first level describes the appearance of the face and hair, the second level describes the refined components of the face (eyes, nose, etc), and the third level defines the wrinkles and creases on the skin.

In [20] an AND-OR graph is also used as a face representation model, but the goal is to reconstruct the input images and generate their caricatures. The graph has three decomposition levels: the first level represents the whole face, the intermediate level decomposes the face into its main parts (eyebrows, eyes, nose, mouth and the rest of the face) and the lowest level divides the parts into small fragments. The model has dictionaries of image primitives for every level, from which the reconstructions and generations of caricatures are constructed.

Another example of the application of AND-OR graphs in facial images proposes a tool that can apply cartoon-like styles in face images [21]. This kind of style is achieved by the intensification of colors, intensification of the contours and attenuation of textures. In this work the faces are represented by a two-level AND-OR graph: the first level has a representation of the whole face that occupies the root of the graph, the second level shows the face divided into eyebrows, eyes, nose, mouth and the rest of the face. A syntax analysis algorithm analyzes an input image in order to identify the face components to identify the edges and protrusions that will be highlighted in the next step.

In short, no examples were found related to the AND-OR graph applications to generate faces without the use of an input image. The most frequent structural applications are recognition, reconstruction and transformation of images.

Furthermore, as observed in [20] and [21], a face AND-OR graph generally has an AND node representing the entire face as root, which immediately decomposes in its main parts. In [20] and in [13], there are graphs that have, in the last decomposition level, nodes that point directly to image primitives, in pixels, already available in libraries. Further on it will be seen that these characteristics are not suitable for the AND-OR graph of this work, and that other approaches were used to replace them.

III. Methodology

The methodology used in this work is divided into two parts: generation of facial emotion caricatures and definition of the AND-OR graph to represent these caricatures. The caricatures were generated using graphics processing concepts presented in section **??** and detailed in section **??**. The measures and parts of faces defined in the graphics processing contains the information needed to define the AND-OR graph described in section **??**.

The caricatures to represent the facial expressions were constructed considering that each part of the face is composed of spline curves. Anthropometric measurements were used to define the size of each curve and the correct position of its control points. To create a neutral face, the number of curves and their positioning were determined based on the descriptions of the Action Units (*FACS*). The *FACS* concepts were also used to reposition the control points of the spline curves in accordance with the emotion, in order to reproduce the desired facial expressions.

The design of the curves representing parts of the face is performed by using *B-splines* curves. Each curve is characterized by the vector of control points that shape it.

Most of the curves were generated with three control points, as shown in Figure 2. It should be noted that two points were added to the eyebrow (*pas1* e *pas2*) so that the movement result would be the same as that described in the AUs 1, 2 and 4 (Figure 1). These two points were determined by the spline subdivision method described in [22], which adds more points to the curve without changing it. The coordinates of the new point are defined as the average coordinates of two points of the curve.

Anthropometric measurements vary with the groups ethnicity and with the individuals age. This work used the data of a population composed of healthy Caucasian North American individuals (19-25 years of age), an age in which adult measurements stabilize [9].

Anthropometric measurements refer to the distances between *landmarks* Therefore, to draw a neutral face required calculating the position of each of these points in the two-dimensional space of the image based on the distance ratio between two points, provided by anthropometric measurements.

Several of the points defined in this work do not have equivalent *landmarks* in anthropometry. To set the position of these points required an approximation based on available anthropometric measurements, while taking into consideration the symmetry of the face. This can be exemplified with the definition of point *p1* (Figure 2). The eyebrow line is positioned horizontally on the same coordinate of point *en*. Besides this point, points *p2* of the eyebrow, *pp1* and *pp2* of the lower eyelid, as well as the wrinkle points, do not have any anthropometric descriptions, they were therefore inferred based on other anthropometric measures in order to represent their curves.

To calculate the positions of the points, a starting point was defined, the *landmarks v*, which is based on the other points. Point *v* is positioned at the top of the head and was defined as the upper end and vertical center of the image (Figure 2).

To show how the position of a given point is computed, note the *sn* point, which is drawn at the bottom part of the nose. Its vertical position was defined as the vertical position of point *n* added to the anthropometric measurement *n-sn* and its horizontal position is equal to the horizontal position of *n*. Thus, all control points of all the curves that form the face are defined.

To reproduce facial expressions in the *sketchs*, the effect of facial emotions was reproduced applying it effect to the

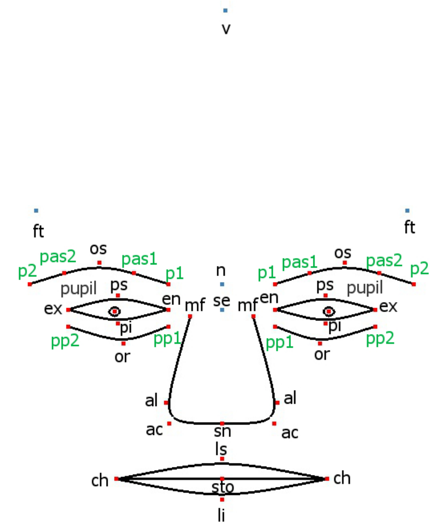


Figure 2. *Landmarks* used for the construction of sketches. The red markings represent the control points used in the curves, and the blue markings are auxiliary points used for positioning the other control points. In addition, green letters describe new points created to build the curves.

curves that compose the face.

The facial emotions to be defined were first represented in this study as they are universal expressions. Then, the action units that shape these facial expressions were studied, as well as the facial appearance resulting from the influence of each unit [11].

No descriptions were found for the varying intensity of the muscles involved in the action units. For example, in action unit 2, FACS refers to the arching of eyebrows, stretching of the upper eyelid and a possible wrinkling of the forehead (Figure 1(b)). However, the degree of arching, the wrinkle measurement and how much the eyelid is stretched are not defined. This problem is bypassed by determining an imaginary rectangle range for each curve. This rectangle range is defined in order to represent a position variation of twice the size of an internal rectangle that encloses the curve tightly, as illustrated in Figure 3. Thus, repositioning a point is based on a percentage of the corresponding measurement range, because the dynamic generation of expressions allows the user to set the emotion percentage that should be represented (from zero to 100% for each type of emotion).

The approach presented enables to generate from more subtle expressions to full expressions from the varying control points involved in each action unit of the facial emotion requested. The transitions between two different facial expressions are achieved passing, through the neutral expression. Examples of different percentages for the same expression are shown in Figure 5.

Since anthropometry does not describe the FACS facial wrinkling, it was necessary to estimate the location of points to generate wrinkles and the variations they undergo. To provide greater realism to some expressions that include

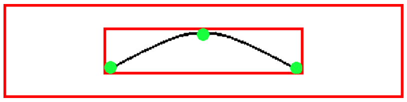


Figure 3. The internal red line shows size of an wrap in relation to the enclosed curve. The outer line delineates a wrap that is twice the size. The black curve corresponds to some part of the face and the green points are its control points.

wrinkling, the wrinkles move according to that part of the facial movement in the corresponding AU. For example, the wrinkles on the forehead have the same vertical elevation as the eyebrow to represent the *surprise* expression.

The AND-OR graph proposed herein, and described below, is only partially sketched in Figure 4 for the sake of clarity. The Figure shows only the face options of *neutral*, *satisfaction* and *sadness*, and at the intermediate levels there are nodes and edges that are not shown. The last levels (curves and points) are only symbolized in the figure for the sake of clarity.

The full definition of the graph presents an OR node at the root (level 1), which selects one of seven different AND nodes, each symbolizing a face with one of the six basic facial emotions, in addition to the neutral face (level 2). The third level has the OR nodes of the seven main parts that make up any of the face types: eyebrows (left and right), eyes (left and right), nose, mouth and the rest of the face. As these are the OR type, they are also selectors and each have a range of options at level 4 (such as straight mouth, upward curved mouth, downward curved mouth, open mouth, etc).

In this design, the goal is to generate images by drawing curves without the use of primitive libraries. The technique for drawing the *b-spline* requires only a set of control points as additional information to draw a curve. To do this, two new classes of components were included in the graph: curves and points. Each AND node at the fourth level decomposes into curves, at the fifth level.

The remaining face area will have the curves that define the face contour, as well as several internal curves that serve as facial expression marks and other details. At this fifth level the curves are OR nodes, and also have types to choose from the sixth level. Finally, each curve in an AND node at the sixth level is composed of the control points used as a parameter to generate them. The points are the leaf nodes of the graph and contain their respective x and y coordinates in the two-dimensional plane.

A given configuration of this graph represents an initial face with a particular emotional expression, but the leaves have a possible positioning range of all control points. Thus, to represent a given image, the exact values of x and y coordinates of the control points must be defined. This is done by randomly initializing these values in a range of one standard deviation to more or less averages (values defined in the leaves based on anthropometric data). These values

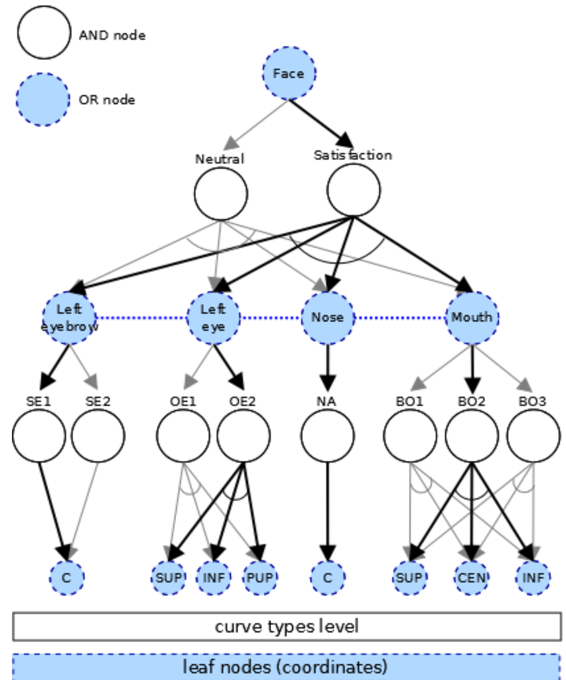


Figure 4. *Illustration of partial AND-OR graph with the neutral face and the happiness emotion.*

are then adjusted according to the desired intensity of that emotion. Therefore, many different images can be generated from a similar configuration.

To generate the image configurations, an algorithm must recursively traverse the graph from the root, extending the search in each AND node and selecting an option from each OR node. During this trajectory, the distance relationships are consulted and the positions of the points are selected and adjusted according to the predefined constraints in the relationships. The black arrows in Figure 4 indicate the edges associated with the productions that create a facial expression of satisfaction.

Most horizontal edges indicate spatial distance relations between the parts, curves and points. These distances, as mentioned earlier, are designated according to anthropometric data.

Volunteers evaluated the realism of the caricatures generated. Forms were constructed using the *GoogleForms* tool, available on the Internet (http://www.google.com/google-d-s/createforms.html). [1].

Each form had four images with different intensity levels (25%, 50%, 75% and 100%) of each of the six emotions and an image representing the neutral face. As a comparison mechanism, the six images with 75% intensity were repeated at random positions in the forms, one from each non-neutral emotion, in order to estimate the evaluators consistency level. Thus, each form comprised 31 images. The evaluator

[1] A demo form is available at: http://goo.gl/UhXwq7.

was asked to choose witch of emotions were represented in the image.

The participation of the evaluators was voluntary and anonymous. The forms were distributed through electronic messages by the project participants. Age and gender data were also collected to allow verifying perception differences according to these variables.

IV. RESULTS

Figure 5 illustrates the part of the caricatures generated for the evaluation process (intensities 25% and 100%). The complete set of images can be seen at the demonstration form cited before.

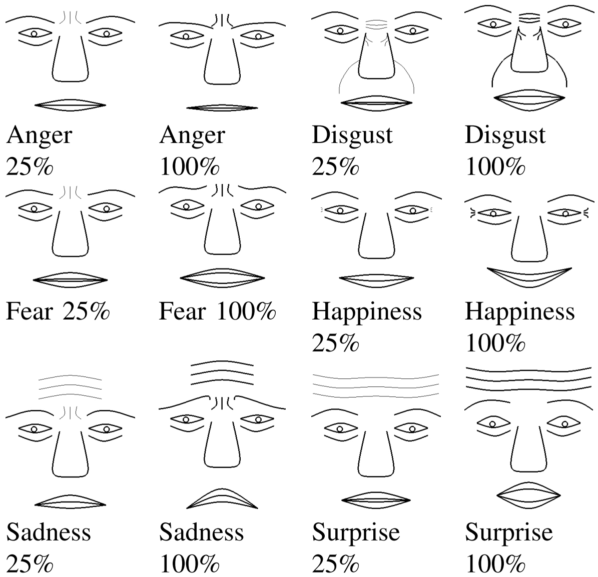


Figure 5. Facial expressions used in the evaluation.

The evaluation of the images included 237 healthy volunteer evaluators, with ages ranging from 13 to 77 years, 54% of all evaluators were men. Three age groups were considered in the results presentation: less than 25, 26 to 55 and greater than 55 years. The percentage of participants in each group was, respectively, 32% , 61%and 6%.

Three age groups were considered in the results presentation: less than 25, 26 to 55 and greater than 55 years. The percentage of participants in each group was, respectively, 32% (77 volunteers), 61% (145 volunteers) and 6% (15 volunteers).

Figure 6 shows each curve that represents the percentage of correct answers for each of the intensities used for each emotion (25%, 50%, 75%, 100%). Table I shows the cross-referencing of responses between the different types of emotions represented in the images. It should be noted that for the 75% intensity of each emotion an average was taken because of duplicate images.

A final analysis was carried out for the responses. As described in the previous section, images were repeatedly included in the forms as a control form. Table II shows the levels of responses of repeated images, considering an overall analysis and divisions by gender and age group.

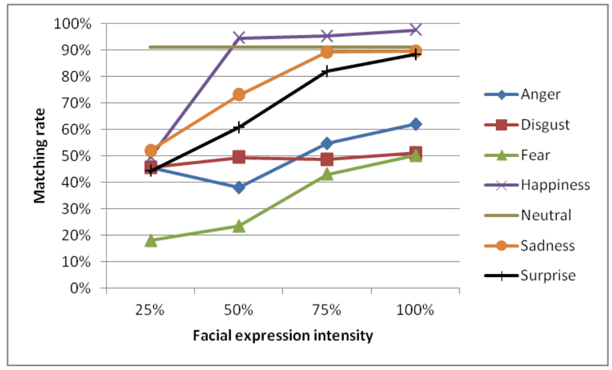


Figure 6. Rate of correct responses for each emotion evaluated.

Table II
CONSISTENCY OF RESPONSES REGARDING THE DUPLICATE IMAGES.

	<25%	**26 a 50%**	**51 a 75%**	**>75%**
Global	1%	28%	33%	38%
Male	2%	28%	29%	41%
Female	1%	28%	37%	34%
<26 years	0%	24%	36%	40%
26-55 years	2%	27%	32%	39%
>55 years	0%	56%	25%	19%

Though different intensities were used for each emotion in the interviews, the correct response considered for an image was the emotion related to that image, though no intensity related questions were asked. It should be noted that the error rates are not indicative of the evaluators inability to recognize facial emotions, but rather methodology deficiencies to be remedied in the representation of emotions in the caricatures.

Figure 6 shows that, except for the neutral face and the *disgust* emotion, the correct recognition rate increased as the intensity of the emotions represented also increased. Another exception regards the *anger* emotion, which at the 50% intensity showed a decrease in the correct response rate, when this emotion was often confused with the *neutral* emotion. The two emotions differ in the downward arching of the mouth and the inclusion of wrinkles (frown) between the eyes in the *anger* emotion, which are subtle at intensities of 25 and 50%.

The images related to the *fear* emotion were those that had the lowest recognition levels in all intensities, followed by images related to *disgust* and *anger* emotions. The analysis of the images representing these emotions shows that these are the caricatures with the least changes in the lower face (mouth region). This suggests that the volunteers can easily recognize the caricatures with more pronounced changes in this region.

Table I
CONFUSION MATRIX OF THE RESPONSES OBTAINED. BOLDFACE INDICATES RECOGNITION RATES OF INTENDED EMOTION.

	Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
Anger	**51%**	7%	6%	1%	27%	5%	3%
Disgust	38%	**49%**	2%	2%	5%	2%	2%
Fear	9%	7%	**36%**	2%	12%	13%	21%
Happiness	0%	0%	0%	**86%**	10%	2%	2%
Neutral	1%	0%	2%	2%	**91%**	2%	2%
Sadness	6%	5%	6%	0%	3%	**79%**	1%
Surprise	0%	1%	8%	5%	13%	2%	**71%**

Table I provides more detailed data regarding the confusion in recognizing the emotions. Analyzing the *fear* emotion on the third line of this table demonstrates that there is a lower recognition rate (36%). The images that are related to this emotion were recognized by the volunteers as belonging to nearly all the other emotions. The emotion in which the *fear* images was most confused regarded the *surprise* emotion.Analyzing the images of these two emotions it was noted that the lower facial sides are similar, increasing the mouth opening according to the intensity. Thus, it evidences that this facial change was related to the *surprise* emotion and that the caricatures related to the *fear* emotion should be reassessed in order to add or change features that can better represent this emotion.

The *disgust* emotion was often confused with the *anger*. This shows that difference nose and forehead wrinkles was not enough to clearly express the aforementioned emotion. Similarly, the users confused the *anger* emotion with the neutral face, indicating that the features of this emotion should be more intense.

With the exception of the three emotions analyzed, the main diagonal of the confusion matrix shown in Table I indicates that all other emotions presented satisfactory recognition rates. The *happiness* emotion presented the highest correct recognition rate, followed by the expressions of *sadness* and *surprise* emotions. These data suggest that the caricatures were properly recognized by the evaluators, especially with intensities of 75 and 100%.

These results agree with the results described in [16], which report that the expressions of *happiness*, *sadness* and *neutral* emotions had high recognition rates, while the *disgust* emotion was vaguely recognized. Dyck and colleagues, however, achieved satisfactory recognition rates for the *fear* and *anger* emotions [16].

On Table II each column shows the percentage of subjects that attained a particular consistency range. A given result was regarded as consistent when the evaluator indicated the same emotion for two identical caricatures in the form, although this emotion was not correct.

The first line of the table indicates that over 70% of the volunteers attained over 50% of consistency. Lines 2 and 3 indicate that there was no significant difference by gender.

However, consistency reduced as the volunteers age increased. In the youngest age group (under 25) 76% of the volunteers showed a higher consistency range. However, the volunteers with ages between 26 and 55 as well as over 55 years showed, respectively, lower consistency averages of 71% and 44%. Although the lower number of volunteers over 55 may have influenced these percentages, it should be highlighted that the older evaluators were more confused with the caricatures than the younger ones.

V. CONCLUSION

The approach used to produce caricatures with different emotional intensities achieved encouraging results with regards to the evaluation conducted.

The caricatures of the neutral face and the expressions of *happiness*, *sadness* and *surprise* emotions were clearly defined by the traces using *spline* curves, however the *fear*, *disgust* and *anger* emotions included caricatures that had a limited recognition range. We believe that refining the features and including new features can reduce the confusion in representing these emotions as well as improving the results.

The importance of undertaking the task to generate virtual images instead of using real images should be emphasized. Virtual images for identifying emotions can be equally effective when compared with real images in research applications [16], [15]. Moreover, virtual images make it is easier to implement applications for the manipulation and animation (dynamic changes between intensities and emotions) of faces, hence enabling greater precision [15]. For example, it is not possible to objectively define how a 75% expression of *disgust* should be made. In addition, a real person may not be able to separate the specific muscles of a particular emotion, which results in mixed expressions, consequently making them more difficult to recognize. The specific goal of this work regards creating a game for training individuals with emotion recognition impairment, and a less complex virtual scenario could represent a preparation process for real-life situations.

VI. ACKNOWLEDGMENTS

This research was supported by the State of São Paulo Research Foundation (FAPESP grant #2011/50761-2, CNPq, CAPES, NAP eScience - PRP - USP), Brazilian National Council of Scientific and Technological Development (Processes 559931/2010-7 and 401745/2013-9 and PIBIC Program), the National Institute of Science and Technology

Medicine Assisted by Scientific Computing and Provostś Office for Graduate Studies of the University of So Paulo. We would like to thank the BIOINFO-Vision Laboratory (University of Sao Paulo) for computing facilities.

REFERENCES

[1] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion." *Journal of personality and social psychology*, vol. 17, no. 2, p. 124, 1971.

[2] C. C. d. A. Rocca, E. v. d. Heuvel, S. C. Caetano, and B. Lafer, "Facial emotion recognition in bipolar disorder: a critical review," *Revista Brasileira de Psiquiatria*, vol. 31, no. 2, pp. 171–180, 2009.

[3] P. Gosselin, G. Kirouac *et al.*, "Decoding facial emotional prototypes," *Canadian journal of experimental psychology*, vol. 49, no. 3, p. 313, 1995.

[4] M. B. Harms, A. Martin, and G. L. Wallace, "Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies," *Neuropsychology review*, vol. 20, no. 3, pp. 290–322, 2010.

[5] S. F. Taylor and A. W. MacDonald, "Brain mapping biomarkers of socio-emotional processing in schizophrenia," *Schizophrenia bulletin*, vol. 38, no. 1, pp. 73–80, 2012.

[6] U. Lahiri, E. Bekele, E. Dohrmann, Z. Warren, and N. Sarkar, "Design of a virtual reality based adaptive response technology for children with autism," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 21, no. 1, pp. 55–64, 2013.

[7] Y. Cheng and S. Ling, "3d animated facial expression and autism in taiwan," in *Advanced Learning Technologies, 2008. ICALT'08. Eighth IEEE International Conference on*. IEEE, 2008, pp. 17–19.

[8] O. Grynszpan, J.-C. Martin, and J. Nadel, "Multimedia interfaces for users with high functioning autism: An empirical investigation," *International Journal of Human-Computer Studies*, vol. 66, no. 8, pp. 628–639, 2008.

[9] L. Farkas, *Anthropometry of the head and face*. Raven Press, 1994.

[10] P. Ekman and W. V. Friesen, "Measuring facial movement," *Environmental Psychology and Nonverbal Behavior*, vol. 1, no. 1, pp. 56–75, 1976.

[11] ——, *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Palo Alto: Consulting Psychologists Press, 1978.

[12] J. D. Foley, *Computer graphics: Principles and practice, in C*. Addison-Wesley Professional, 1996, vol. 12110.

[13] L. Lin, T. Wu, J. Porway, and Z. Xu, "A stochastic graph grammar for compositional object representation and recognition," *Pattern Recogn.*, vol. 42, no. 7, pp. 1297–1307, Jul. 2009.

[14] S.-C. Zhu and D. Mumford, *A stochastic grammar of images*. Now Publishers Inc, 2007.

[15] E. Bekele, J. Crittendon, Z. Zheng, A. Swanson, A. Weitlauf, Z. Warren, and N. Sarkar, "Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism," *Journal of Autism and Developmental Disorders*, Jan 2014.

[16] M. Dyck, M. Winbeck, S. Leiberg, Y. Chen, R. C. Gur, and K. Mathiak, "Recognition profile of emotions in natural and virtual faces," *PLOS One*, vol. 3, no. 11, p. e3628, 2008.

[17] J. Ku, H. J. Jang, K. U. Kim, J. H. Kim, S. H. Park, J. H. Lee, J. J. Kim, I. Y. Kim, and S. I. Kim, "Experimental results of affective valence and arousal to avatars facial expressions," *Cyber Psychology & Behavior*, vol. 8, no. 5, pp. 493–503, 2005.

[18] D. DeCarlo, D. Metaxas, and M. Stone, "An anthropometric face model using variational techniques," in *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '98. New York, NY, USA: ACM, 1998, pp. 67–74.

[19] J. Suo, S.-C. Zhu, S. Shan, and X. Chen, "A compositional and dynamic model for face aging," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 3, pp. 385–401, March 2010.

[20] Z. Xu, H. Chen, S.-C. Zhu, and J. Luo, "A hierarchical compositional model for face representation and sketching," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 6, pp. 955–969, June 2008.

[21] M. Yang, S. Lin, P. Luo, L. Lin, and H. Chao, "Semantics-driven portrait cartoon stylization," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, Sept 2010, pp. 1805–1808.

[22] D. Hearn and M. P. Baker, *Computer graphics, C version*. Prentice Hall Upper Saddle River, 1997, vol. 2.