Hidden Emotions in Iconic Faces

Alicia Heraz

DIRO. University of Montréal Pavillon André-Aisenstadt. 2920, chemin de la tour alicia.heraz@umontreal.ca

Claude Frasson

DIRO. University of Montréal Pavillon André-Aisenstadt. 2920, chemin de la tour claude frasson@umontreal.ca

Abstract. Most of the time, humans use texts to exchange information and more recently they can add iconic faces (or smiley) to introduce emotional content. What emotions can be conveyed by iconic faces? To understand better this question we conducted an experimentation in which we asked participants to interpret a set of iconic faces by selecting an emotion in a proposed list of faces. Results indicated that some unexpected emotions, such like deception or desire were identified on iconic faces. That is why we believe that the use of iconic faces can convey more emotions and consequently should improve human computer interaction. We proposed also to use machine learning techniques to recognize and classify emotions. We used a specific method to train a Support Vector Machine and we reached 67% of accuracy.

Keywords: Iconic Face, Facial Expression, Emotional Recognition, Machine Learning, Support Vector Machine.

INTRODUCTION

During the last years, we have seen a large increase in research directed towards affective dimension of human computer interaction. In recent Intelligent Tutoring Systems, for instance, learners can communicate via forums, chats or emails to cooperate. In fact, these tools become a central communication channel for private and educational exchange. Information exchanged between learners may concern courses, exercises, difficulties in comprehension, but also current feelings, emotions and moods. However, we may be unable to recognize facial expressions on some learner faces, because of impassive faces, cultural barriers or timidity. If we could attach emotions to a text, communications between human could be enriched by human factors. Including iconic faces, for instance, can be worth more than words. But which emotions are conveyed by which faces? In most popular forums and instant messaging servers, humans can use a finite set of iconic faces, to express a restrictive set of emotions.

In this paper we want to explore different parameters related to emotions on iconic faces and find what emotions can be recognized on that faces. How people interpret iconic faces and what they recognize as emotions? For the first part of the question we asked 354 participants to interpret randomly generated iconic facial expressions [Chernoff, H., & Rizvi, M. H. 1975] and for the second part we used a Machine Learning technique to automatically provide iconic facial expression recognition. Computers may never need all of the emotional skills that people need; however, it is obvious that computers will require at least some of them to appear intelligent when interacting with people [Picard, Healey and Vyzas 2001]. Recognition of emotions is part of emotional intelligence [Goleman D., 1995]. But what does an iconic face convey? Eckman identified six basic emotions on human facial expression: joy, sadness, fear, anger, disgust and surprise [Eckman P., 2002]. Are these emotions similarly recognized using iconic faces? This paper argues that some emotions can be more clearly recognized on some iconic faces than on human facial expression. The discussion below is divided into three principal sections: state of the art (section 1 and 2), data collection analysis (section 3) and classification results (section 4). In the first section we address different ways to express emotions, whereas in the second section we demonstrate the utility of iconic faces in Intelligent Tutoring Systems; in the third section we detail the experimentation principles and analyze the results of iconic faces interpretations; finally, in the fourth section we give classification results and improvements made on data collection to minimize error of classification and maximize accuracy of the recognition system.

1. EMOTIONS EXPRESSION AND RECOGNITION

In this section, we analyze the process of communication between learners using text mixed with iconic expressions and the possible difference that can exist between expression and interpretation of emotions.

1.1. Emotions and words

Writing is probably the largest and simplest mean for remote communication via computers, but communication of emotional expression in writing requires considerable skill. Generally people use different words to convey emotions but they may interpret differently the emotional degree of these words. This is why many research groups collected and classified huge number of words which might be emotional words. [Shields, 1984] collected 60 words, [Storm and Storm, 1987] produced a list of 72 words and they extended their list to 193 terms resulting from several hundred of queries asked to children and adult. Recently, Philipp Shaver and his colleagues based on research they conducted with college students [Parrott, W. 2001]. They rated 213 emotional terms on how much emotion they conveyed. On the basis of students' descriptions, the researchers converged to 5 or 6 clusters that they believed represent basic emotional categories. In the present paper we will use the taxonomy of Philipp Shaver (Figure 1) which identifies three levels of emotions: primary, secondary and tertiary emotions.

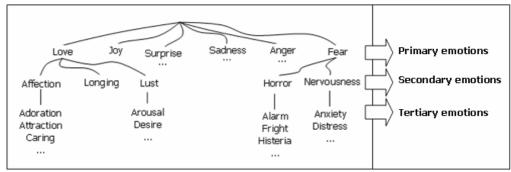


Figure 1: tree like taxonomy of Philipp Shaver

1.2. Emotions and faces

We consider emotion as important information that we need to transmit to enrich a message. Though some emotions can be perceived in tone voice [Scherer K. R. and Oshinsky J. S., 1977; Ohala, J. J., 1996], most of the emotions are perceived through facial expressions. Eckman and al. argued that even impassive faces always transmit emotions.

People can identify basic emotions from facial expressions with high accuracy [Ekman, P., 2002]. The faces help to moderate and choreograph conversations [Donath, J., 2001]. The face conveys information through its structure, its dynamics, and its decorations [Zebrowitz, L., 1997]. Even if all emotions are not easily and clearly perceived on a human face, we believe that each emotion has a specific facial expression with flagrant or subtle modifications. For example, for sad, mournful expressions we use eyebrow depression, eye narrowing, nasal muscle elevation, lip compression, and mouth lowering (turned down at edges).

1.3. Human faces or Iconic faces?

If I am happy, I open mouth, bridle my eyes and decline my eyebrow, but what's happened on my face when I feel desire, deception, anxiety or timidity? These examples of emotions are not identified by Eckman studies on emotions and facial expressions. We think that what human can perceive through other human faces is limited and can differ from the reality. Many studies denote a variation in subtle parameters on human face such as pupil size, but these parameters are not explicitly seen by a human on a human face.

In an iconic face, we can exaggerate these parameters to voluntarily express some kind of emotions; iconic face is an extensive caricatured face. By emphasizing on certain subtle parameters on iconic face, we can convey more information, more than the six basic emotions detected on a human face by Eckman, which are: joy, sadness, surprise, disgust, fear and anger [Eckman P. 2002]. An example of a parameter which is not easily perceived on a human face is the pupil size. Pupils show a desire to improve and achieve in relation to their own abilities. Nonverbal communication is part of the communication which is not based on words or phrases. They include gestures, attitudes, movements of eyes, pupil size, lip thickness, skin complexion and they represent two thirds of what human express unconsciously.

Pupil response is a nonverbal expression of strong emotion. Enlarged pupils mean acceptance, warmth, sexual attraction, or strong surprise. Very small pupils can indicate dislike or a rejection of what is being viewed. Of course, pupils can change even if we are thinking about something which is not necessarily present in the real

life. Most small children are aware of the significance of pupil size; all of the Walt Disney cartoons use pupil size to designate who is mean and whom is to be trusted.

When we see a cartoon that has a "flat" feel, the artists have not gone to the trouble to use pupil size to indicate who is the hero or the heroine. Facial expressions of the primary emotions are not learned; for instance, children born deaf and blind can display all of the correct expressions for sadness, happiness, and anger [Psychology department San Antonio College. 2005] without any learning. It was found that caricatured faces were judged as more clearly representing each emotion than were the composite faces. It was also reported that there were confusion errors between fear and surprise and, to a lesser degree, between sadness and anger. The caricatured images also did not speed up the naming of the facial expressions. This may have been due to the possibility that caricature morphing increased the appearance of mixed emotions on the face [Plutchik R. 2002].

2. ICONIC FACES, VIRTUAL COMMUNICATION AND EDUCATION

Most of Intelligent Tutoring Systems are equipped with online communication tools such as chat and forum [Silveira and Vicari, 2002] in order to provide students with more resources to solve doubts and learn more about topics. An iconic face can be illustrated by a smiley during student communication. In addition, iconic faces can be visual representation of pedagogical agents or personified agents. The use of personified agents could be beneficial for human-computer interactions [Laurel, 1997] expressing emotions through iconic facial expressions. [Rousseau and Hayes-Roth, 1998] developed an extensive social-psychological model that considers agent personality, emotions, and attitudes so that the agent can select an appropriate behaviour to enhance its believability. [Gratch, 2000] created a "personality GUI" that contains the agent's goals, social status, etiquette, (in)-dependence and attitudes towards the user and other agents.

Importantly, there is significant empirical evidence that by rendering systems more human-like users can rely Agent Persona 3 [Baylor, & Ryu, 2003] on standard interaction skills (such as interpreting agents' facial expressions or taking into account eye contact), making the interaction with the computer much smoother [Dehn & van Mulken, 2000]. The problem is then expressed by the following questions: Does the learner perceive on iconic faces the same emotions which he perceives on human faces? What emotions can be expressed by an iconic face?

3. HOW HUMANS INTERPRET ICONIC FACES?

To understand better how this interpretation is effectively performed, we have developed a data acquisition tool. The goal of the project is to analyze different facial interpretations and also to constitute a training data set as input to a machine learning system.

3.1. Description of the experimentation

We have developed and deployed on the Web an interface in which we asked people from different countries to interpret randomly generated iconic facial expressions by associating one of the Shaver's 213 emotions organized according to the hierarchy described in section 1.1.

Experiments were realized in two steps: in the first step, participants try to recognize a facial expression within one of the six primary emotions; in a second step, they can precise more their perception by choosing another emotion within the secondary or a tertiary list of emotions derived from the primary emotion. In addition, each participant has to complete information about his profile with some personal and cultural parameters such as sex, age, country... for further data analysis. Over 789 participants produced 31,986 interpretations.

3.2. Results and discussion

3.2.1. Ambiguity and confusion

The first thing that we can conclude is that there is confusion in the interpretation of emotions. An iconic face may represent sadness for one, while it may represent fear for another one. Figure 2 gives a summary on most of the confusions. We observed that people have confusion between emotions such as love and joy, fear and sadness and also surprise and sadness.



Figure 2: confusion between emotional iconic faces

Neutral and undefined faces were also sometimes confused with sadness, joy, anger or surprise and almost never confused with love and fear.

3.2.2. Frequently perceived emotions

Figure 3 gives the recognition percentages of emotions that are generally well recognized. Sadness, for instance, is frequently identified on an iconic face.

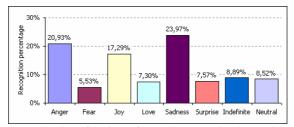


Figure 3: classification of mostly recognized emotions

3.2.3. Response time

To recognize an iconic face, participants have taken statistically an average of 20.87 seconds to decide if they see an emotion, 16.45 seconds if they perceive an ambiguous face and 9.84 seconds if the face is neutral. Figure 4 shows the repartition of the average time taken by participants to interpret an iconic face. We observe that a neutral iconic face is more quickly recognized than the other emotional faces.

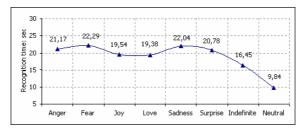


Figure 4: Average time required to identify an iconic face

3.2.4. Primary, secondary or tertiary emotion?

Participants had to successively choose from a hierarchical list of emotions, between primary, secondary and tertiary emotions. We observed that more the emotions are precise more they can be recognized. Figure 5 shows for each emotion, the percentage of recognition between primary, secondary and tertiary emotion. Tertiary emotions are more frequently recognized, especially for love and fear, and this may be due to the fact that fear and love are complicated terms; derived emotions which are more precise reduce this ambiguity.

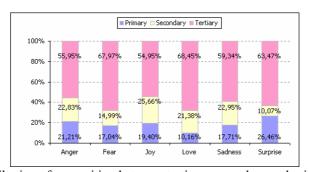


Figure 5: distribution of recognition between tertiary, secondary and primary emotions

For the primary emotion love, secondary and tertiary derived emotions are classified in the ten mostly recognized emotions (for example: affection, desire and adoration). Table 1 gives the top 10 frequently recognized emotions.

Emotion	Percentage	Type
Astonishment	% 26.90	Tertiary, from Surprise
Amazement	% 15.49	Tertiary, from surprise
Disappointment	% 7.92	Tertiary, from sadness
Surprise	% 7.64	Primary
Affection	% 6.43	Secondary from love
Desire	% 5.22	Tertiary from love
Adoration	% 4.09	Tertiary from love
Anxiety	% 3.55	Tertiary from fear
Carinq	% 3.06	Tertiary from love
Contentment	% 2.91	Secondary from joy

Table 1: top ten of mostly recognized emotions

Surprisingly, we also denoted recognition of some unexpected emotions, such as desire and disappointment.

4. USE OF MACHINE LEARNING FOR RECOGNIZING ICONIC FACES

Using now the results obtained from the previous data collection, we thought to insert these results into a Machine Learning algorithm in order to automatically recognize emotions. Then, a large number of recognitions could be taken into account to deduce how people can generally interpret iconic expressions.

The learning method that we decided to choose was the Sequential Minimal Optimization (or SMO) because it is a fast method to train Support Vector Machines (SVMs). Training a SVM requires the solution of a very large quadratic programming (QP) optimization problem. SMO breaks this large QP problem into a series of smallest possible QP problems. They are solved analytically, which avoids using a time-consuming numerical QP optimization as an inner loop [Cristianini, N., Shawe-Taylor, J. (2000)].

4.1. Machine learning model

Determining the emotion conveyed by an iconic face can be cast as a multi-class classification problem. The classification problem is defined as follows: Let Cf representing the iconic face generated by Chernoff algorithm, and p_i where $i \in [0..9]$ the ten parameters $(p_0, p_1, ..., p_9)$ which define the characteristics of a Chernoff-like face. Let k be the number of emotion classes $E = \{e_1, e_2, ..., e_k\}$, where e_1 denotes the special case of neutrality or absence of emotion and e_2 denotes an unrecognized emotion in case of ambiguous iconic face. The mapping function to find is $f: Cf(p_0, p_1, ..., p_9) \rightarrow e_j$ where $j \in [1..k]$, such that we obtain an ordered labelled pair $(Cf(p_0, p_1, ..., p_9), e_j)$. Even if our first goal is to predict finer emotional meaning according to the parameters, we focus here on acceptability of error classification, despite the huge number of participants in this experimentation. We need to know if we have too much divergence among the interpretations.

4.2. Classification results

For classification we used Weka, a collection of machine learning algorithms intended for data mining. The algorithms can be either applied directly to a dataset, or called from a Java program. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes [Ian H. Witten and Eibe Frank (2005)]. In the experimentation we fixed the complexity parameter c, used in the SMO function, to 100.0 and the gamma parameter to 0.05. 70% of data were used for training, and 30% for test (31,986). Despite this small test data set, we obtained promising results. Figure 6 shows the evolution of accuracy with the size of data considered. It seems, finally, that we can obtain higher accuracy and convergent opinion with iconic facial expression of emotions. Table 2 shows a summary of global results and table 3 shows the details of the accuracy.

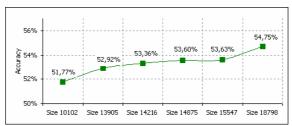


Figure 6: Evolution of accuracy among set sizes

Correctly Classified Instances	% 54.75
Incorrectly classified Instances	% 45.25
Kappa statistic	0.4288
Mean absolute error	0.1978
Root mean squared error	0.3083
Relative absolute error	% 94.74
Root relative squared error	% 95.45

Table 2: summary of the evaluation on test split

TP Rate	FP Rate	Precision	Recall	F-Mesure	Class
0.008	0.002	0.2	0.008	0.015	Love
0.897	0.183	0.513	0.897	0.653	Joy
0.135	0.022	0.343	0.135	0.194	Surprise
0.785	0.149	0.606	0.785	0.684	Anger
0.813	0.201	0.562	0.813	0.665	Sadness
0	0	0	0	0	Fear
0.081	0.012	0.33	0.081	0.129	Neutral
0	0.001	0	0	0	Undefined

Table 3: detailed accuracy by class

Confusions shown in figure 1 are confirmed by confusion matrix produced at the issue of the training. In fact, a lot of faces classified as representing love are classified as representing joy. Sadness is also confused either with fear or surprise. The worst classification is of fear; it seems very difficult to identify parameters that represents really fear, at most of the time, sadness is the last output. Undefined faces are also not classified with high accuracy.

L	J	S	А	D	F	N	U	
0.007	0.774	0.007	0.101	0.101	0	0.010	0	Love [L]
0.005	0.897	0.017	0.049	0.021	0	0.009	0.001	Joy [J]
0.002	0.207	0.135	0.241	0.381	0	0.027	0.007	Surprise [S]
0	0.239	0.041	0.785	0.309	0	0.029	0	Anger [A]
0.002	0.045	0.015	0.117	0.815	0	0.007	0	Sadness [D]
0	0.075	0.041	0.120	0.755	0	009	0	Fear [F]
0.005	0.291	0.055	0.281	0.283	0	0.081	0.005	Neutral [N]
0.002	0.308	0.049	0.314	0.301	0	0.025	0	Undefined [U]

Table 4: confusion matrix

Colorized cases show the higher percentages of classification for each emotion. The best classifications are for joy, sadness and anger. Most of the expressions of love and fear are classified respectively as joy and sadness. Surprise, neutral and indefinite expressions are classified either as joy, anger or sadness.

4.3. Features experiments

Emotions on iconic faces are poorly understood. It is not clear which attributes are more important than others for facial expression recognition. Thus, we grouped attributes in features group and we experimented with different features configurations. Starting with all features, again using 10-fold cross-validation, we reduced progressively the number of feature groups. Feature groups are listed in table 5. Figure 7 shows the accuracy at each step of the cumulative reduction process.

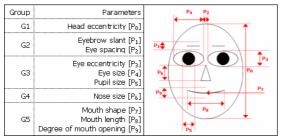


Table 5: feature group members

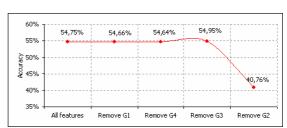


Figure 7: average effect of feature group removal

While some feature groups, e.g. head eccentricity and nose size, appeared less important, the removal order mattered; e.g. if head eccentricity and nose size are removed first, accuracy decreased. This fact also illustrated that features work together, removing any group degraded performance because features interact and there is no true independence.

4.4. Refining the model

We had a fairly small data set to obtain the accuracy of iconic facial expression recognition. Nevertheless, results indicated that our learning approach performs emotion recognition on iconic facial expression. For example, among 789 participants, many identical iconic faces were classified per user in the same outputs; these faces are also correctly classified with high accuracy in the output (with values from 1 to 8). Emotion recognition is part of emotional intelligence; we wanted to see classification results per participant: various selection criteria have been compared empirically in a series of experiments [Quinlan 1983]. At first time, we choose the ten first participants who interpreted the most number of faces; then we constituted ten separated sets and on each one we evaluated accuracy of learning. Results have shown that accuracies for each participant-set were different from global accuracy. Different people can have different emotional reactions to the same stimulus, and the variability depends upon traits that we cannot always easily perceive, such as person's goals, preferences, expectations and personalities. Emotions can be recognized because they often have observable effects on user's behavior and bodily expressions. But the mapping between emotions and their observable effects also depends on often hidden traits of a human face, as well as on the context of the interaction [Conati 2002]. Table 6 shows ten different results of classification per participant. We so think that participants were not able at all to recognize emotions from faces, which means that participants were not equally emotionally intelligent. Accuracy varies from about 50% to 70%. By choosing entries of the first top accuracies, global accuracy of the system increase. We started by considering all participants, step by step, and the lowest accuracy were removed. Figure 8 shows the accuracy at each step of the cumulative substraction process.

Participant ID	Sex	Accuracy
H_0848	М	% 69.41
Y_0917	М	% 68.84
R_1052	М	% 67.41
K_2579	F	% 67.31
M_0503	F	% 63.58
A_1001	F	% 60.80
S_0522	М	% 58.60
A_0552	М	% 56.63
0_1150	М	% 55.07
S_1623	М	% 51.12

Table 6: top ten of accuracy by participant

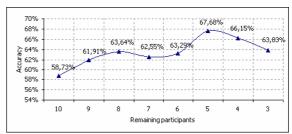


Figure 8: Averaged effect of participant group

Figure 8 shows that to increase accuracy we have to take into account the best participants who obtained the highest accuracy. At the same time, the size of training set must be high, that is why accuracy decreases when only four participants remained.

CONCLUSION AND FUTURE WORK

This paper argues for the use of iconic faces to express emotions during communications between learners and an Intelligent Tutoring System. The first experiment allowed us to understand how participants could interpret iconic facial expressions. Results have shown in fact that they perceived additional emotions, other than those found by Ekman on human faces. We realized that iconic face can be modified to underline a more specific emotion, which cannot be obtained with a human face. The second part of this work presented a classification resulting from a machine learning method able to automatically recognize iconic facial expression.

Results are promising; we reached 67% of accuracy despite the size of training data set. After manipulation and filtration of the training set, we found that to increase accuracy we have to take into account participants with the best accuracy on their entries. Automatic iconic face recognition should improve human computer interactions with an ITS by allowing a machine to understand emotions and consequently provide adapted answers. Presently, due to the lack of data, we did not train our system to recognize deeper emotions like desire or disappointment, so the next step is to do the analysis with more data. We intend to see the evolution of accuracy for detecting more precise emotions.

ACKNOWLEDGMENTS

We address our thanks to FQRSC (Fond Québécois de la Recherche sur la Société et la Culture) and the Ministry of Research, Sciences, and Technology of Quebec for supporting this project.

REFERENCES

- Baylor, A. L. (2000). Beyond butlers: Intelligent agents as mentors. Journal of Educational Computing Research, 22(4), 373-382.
- Baylor, A. L. (2001). Permutations of control: Cognitive guidelines for agentbased learning environments. Journal of Interactive Learning Research, 12(4), 403-425.
- Baylor, A. L. & Ryu, J. (2003). Does the presence of image and animation enhance pedagogical agent persona? Journal of Educational Computing Research, 28(4), 373-395.
- Chernoff, H., & Rizvi, M. H. (1975). Effect on classification error or random permutations of features in representing multivariate data by faces. Journal of American Statistical Association, 70, 548-554.
- Conati, C. (2002). Probabilistic Assessment of User's Emotions in Educational Games. Journal of Applied Artificial Intelligence, Vol. 16, 555-575.
- Cristianini, N., Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines. Cambridge University Press.
- Dehn, D. M., & van Mulken, S. (2000). The impact of animated interface agents: A review of empirical research. International Journal of Human-Computer Studies, 52(1), 1-22.
- Donath, J., (2001) Mediated Faces. In M. Beynon, C.L. Nehaniv, K. Dautenhahn (Eds.). Cognitive Technology: Instruments of Mind. Proceedings of the 4th International Conference, Warwick, UK.
- Ekman, P., Friesen, W.V., & Hager, J.C. (2002) The facial action coding system, Second edition. Salt Lake City: Research Nexus eBook. London: Weidenfeld & Nicolson (world).
- Goleman D. (1995) Emotional Intelligence: Why it can matter more than IQ. New York: Bantam Books.
- Gratch, J. (2000). Socially situated planning. Paper presented at the AAAI Fall Symposium on Socially Intelligent Agents (Agents 2000).
- Laurel, B. (1997). Interface agents: Metaphors with character. In J. M. Bradshaw (Ed.), Software Agents (pp. 67-78). Menlo Park, CA: MIT Press.
- Ohala, J. J. (1996) Ethological theorie and the expression of emotion in the voice. ICSLP 96.
- Parrott, W. (2001) Emotions in Social Psychology, Psychology Press, Philadelphia
- Picard, R. W., Healey, J., Vyzas, E. (2001) Toward Machine Emotional Intelligence Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, p. 1175-1191.(2001)
- Plutchik R. (2002) Emotions and Life, Perspective From Psychology, Biology and Evolution. American Psychological Association, DC.
- Psychology department San Antonio College. http://www.accd.edu/. Motivation and Emotion.
- Quinlan, J.R. (1983) Learning efficient classification procedures and their application to chess end games. In Machine Learning, An Artificial Intelligence Approach, Pages 463-482, Michalski, R. S., Carbonell, J. G., and Mitchell, T. M., Editors, Tioga Publishing Company, Palo Alto, CA, (1983).
- Rousseau, D., & Hayes-Roth, B. (1998). A social-psychological model for synthetic actors. Paper presented at the 2nd International Conference on Autonomous Agents (Agents '98).
- Scherer K. R. and Oshinsky J. S. (1977) Cue utilization in emotion attribution from auditory stimuli. Motivation and Emotion, vol. 1, no. 4, pp. 331-346.
- Shields S.A. (1984). Reports of bodily change in anxiety, sadness and anger. Motivation and Emotion, 8, 1-21.
- Silveira Ricardo Azambuja and Vicari Rosa Maria (2002). Developing distributed intelligent learning Environment with JADE Java Agents for Distance Education Framework. In: Lecture Notes in Computer Science, Volume 2363, Jan 2002,. Berlin: Springer-Verlag, p. 105-118. ISSN 0302-9743
- Storm, Christine; Storm, Tom (1987) A Taxonomic Study of the Vocabulary of Emotions. Journal of Personality and Social Psychology, v. 53: p. 805-816.
- Witten I.H and Frank E. (2005). Data Mining: Practical machine learning tools and techniques, 2nd Edition. Morgan Kaufmann, San Francisco.
- Zebrowitz, L. (1997) Reading Faces. Westview Press, Boulder, CO.