Computer Vision Research at the Computational Vision Laboratory of the Universidad de Chile

Javier Ruiz-del-Solar

Department of Electrical Engineering, Universidad de Chile
jruizd@ing.uchile.cl

Abstract. The development of research in computer vision in a developing country is a challenging task. Factors such as low research funds, low trust from local companies and governments, and small number of qualified researchers complicates the development of strong, local research groups. In this article we present our research group in computer vision at the Universidad de Chile, as an example of how these difficulties can be overcome through international cooperation, focusing on global research projects with low requirements of expensive equipments, formation of qualified researchers, synergistic research with other research fields (e.g mobile robotics), and some necessary basic support from government research agencies. We show how we have been able to contribute in research areas such as face analysis for human-computer interfaces, robust object recognition, and robot vision, at an international level.

Keywords: Human-Computer interfaces, Face Analysis.

1 Introduction

Computer vision is a research field that involves both science and technology development. As in other fields, to carry out research in computer vision in a developing country is a difficult task due to several different factors, mainly:

- *Industry and government support*: In several developing countries the industry and even the government do not trust local research groups for solving their technological problems. They prefer to find technological solutions in institutions from developed countries. This situation frustrates the development of strong, local research groups.

- *Research funds*: Research institutions in developing countries have much lower funds for research than their counterparts in developed countries. The main impact of this issue is the difficult for building long-term research groups. Very often the research is developed only at the universities, and the main researchers are professors working with undergraduates or master students in their projects. In this context it is very difficult to address large or long-term research projects.

- *Availability of qualified researchers*: The availability of specialist in the specific fields (i.e. computer vision) is very low, which also hinder the development of strong, local research groups. Very often the specialists in the field emigrate or they do not return to their countries after a Ph.D. or a specialization, because they find better research opportunities in the first world.

- *Distance to Research Centers and Conference Venues*: Most important research centers are located in developed countries, and most important conferences take place also in developed countries. A consequence of the long distances from developing countries to these centers is the fact that researchers from developing countries have many difficulties for participating in high quality conferences and for visiting top research centers. In addition, a consequence of the low density of researchers in most specific research fields, in developing regions, is the low quality and impact of the regional conferences in these fields.

All mentioned factors are interrelated and it seems we are in the case of a typical *chicken-and-egg problem* situation. We could ask ourselves, which comes first, financial support, trust from the local industry and government, or the availability of qualified researchers? Obviously the answer depends on the country characteristics (GDP – Gross Domestic Product, specific government programs for R&D, industry's innovation culture, etc.), on the research field, and on the specific research groups. We will present as a case of study our research group in computer vision at the Universidad de Chile.

Formally speaking, the *Computational Vision Laboratory* at the Department of Electrical Engineering of the University of Chile was created in 2001. Our main research interest areas are the development of new computational vision paradigms as well as the application of image processing, machine learning and computational intelligence techniques to the resolution of computer vision and pattern recognition problems. Our research work in the last years has been successful in terms of scientific production and projects. We believe that the main reasons are the following:

- *The nature of the research field*: Computer vision is a field in which one can carry out research at the highest level, without requiring high investment in equipments. Of course there are some specific research areas in which expensive equipments are required, but we have concentrated ourselves on the development of object recognition algorithms and human-computer interfaces, which does not require that kind of equipments.

- *International cooperation*: The contact with researchers from other countries helps us in: (i) the development of long-term projects, (ii) the training of students in specific research topics, and (iii) the development of broader projects that require research in several different specific topics. As an example of this kind of cooperation we can mention our long-term cooperation with Dr. Mario Köppen (Kyushu Inst. of Tech.), Dr. Katrin Franke (Norwegian Information Security Lab), and Dr. Aureli Soria-Frisch. (Univ. Pompeu Fabra) in biometry and in the application of soft-computing technologies to computer vision problems. In addition we are currently involve in the ALFA [39] project "Computer Vision Foundations and Applications" with 4 European universities (Ecole Normale Supérieure de Cachan, Univ. Pompeu Fabra, Univ. di Bologna, Univ. de Las Palmas de Gran Canaria) and 2 other South American universities (Univ. de la República, Univ. de Buenos Aires).

- *The global aspects of the applications*: Although in our country there are many potential computer vision applications related with our main industries (e.g. mining, forestry, fish industry), it has been very difficult for us to get funds to develop applied research in those areas. Surprisingly, it has been much easier to develop research projects with international partners. As an example, together with the Fraunhofer Institute IPK [40] we developed in 2005 a face analysis toolbox for the automatic quality control of passport's photographs [41]. Currently, this system is been used in Germany by one of the main companies in charge of passports' production.

- *Formation of qualified researchers*: All the researchers (undergraduates, MS and Ph.D. students, and engineers) that work in our projects have been formed by ourselves. In our department we have one course in image processing, one in computer vision, and several other in related research fields (mobile robotics, neural networks, evolutionary computation, fuzzy systems, control theory, agents, etc.). Thus, in the last 8 years more than 50 students have developed their final engineering or MS thesis (6 years study programs in each case) in computer vision projects, and currently 7 Ph.D. students are carrying out their thesis in the field (our Ph.D. program started in 2005). Most of the students we have already trained in the field are working in computer vision or image processing projects in the Chilean industry. This fact will foster the development of applied research projects in computer vision in Chile. We believe that one of the main achievements of our research group has been the formation of research capabilities in computer vision in Chile.

- *Government support*: The availability of funds for performing basic research activities (i.e. to participate in international conferences, to buy basic computational and sensor equipments) is a necessary condition for carrying out research activities at international level. The Chilean research programs FONDECYT and Millennium have supported our basic research in computer vision, and have gave us the opportunity of building up our research group. In addition, in Chile the research agency CONICYT funds directly Ph.D. students with fellowships. The main requirement for obtaining this support is the quality of the candidates. All our Ph.D. students have these fellowships, which helps us largely in our efforts for building a research group.

- *Synergy with other research efforts*: The UChile robotics team is an effort of the Department of Electrical Engineering of the University of Chile in order to foster research in robotics [1]. The main motivation of the team is the participation in international robotics contests (e.g. RoboCup, ICRA Competition) that provide standard problems to be solved, where a wide range of technologies can be integrated and examined. One of the main research emphases of the UChile team is robot vision. In the last RoboCup competition the UChile team was able to classify their 3 teams of robots in the *four-legged*, *humanoid* and *@home* RoboCup leagues. The research carried out by the UChile team is synergistic with our research in computer vision. Many of the students work in both groups. The application of computer vision to robotics has been a way of attracting engineering students to the computer vision field, and also students to our Ph.D. program.

Since 2001 our contributions to the computer vision field have been mainly in 3 research areas (most of this research has been developed or at least supported by Ph.D. students in their research work):

- *Face Analysis for Human-Computer Interfaces*: We have developed algorithms for face detection [12][16], eyes detection [17], face recognition [13][9], people tracking [20][22], gender classification [5], and skin segmentation [8][18][19] that we have employed in several applications including services robots [27].

- *Robust Object Recognition based on wide baseline matching*: We have developed a robust object recognition system based on SIFTs [7][4], which uses several hypothesis rejection tests. We have applied this system to the recognition of objects in cluttered environment [26], to the detection of articulated robots [7], and to biometric applications [28].

- *Robot Vision*: We have developed several functionalities for our mobile robots, which includes: robot detection using boosted classifiers [23], automatic on-line color segmentation [25], spatiotemporal context-based filtering of perceptions [14][21], gaze direction determination of robots [6][7], motion analysis with moving cameras [10], visual object recognition using genetic algorithms [11], and an automated refereeing and analysis tool for robot soccer [3][15].

Our main achievements in applied research are: (i) the development of a face analysis toolbox for the automatic quality control of passport's photographs, currently in use by a large German company, and (ii) the obtainment of the *RoboCup @Home 2007 Innovation Award* for the human – robot interaction and object recognition abilities of our service robot.

For this article we choose to present some of our research on human – computer interfaces. Thus, in section 2 the main aspects of our face analysis toolbox are described. In section 3 we present Bender, a service robot that uses this face analysis toolbox. Finally, in section 4 conclusions of this work are given.

2 Face Analysis for Human Computer Interfaces

Face analysis plays an important role for building human-computer interfaces that allow humans to interact with computational systems in a natural way. Face information is by far, the most used visual cue employed by humans. There is evidence of specialized processing units for face analysis in our visual system. Faces allow us the localization and identification of other humans, and the interaction and visual communication with them. Therefore, if we want that humans can interact with machines with the same efficiency, diversity and complexity used in the human-human interaction, then face analysis should be extensively employed in the construction of human-computer interfaces.

Face detection is a key step in almost any computational task related with the analysis of faces in digital images. Moreover, in many different situations face detection is the only way to detect persons in a given scene. Knowing if there is a person present on the image (or video) is an important clue about the content of the image. In the case of human computer interaction applications, clues about the gender, age, race, emotional state or identity of the persons give important context information. When having this kind of information, the application can be designed to respond in a different way depending on the user's identity, mood, gender or age. Face recognition systems can be improved by using other clues about the face or by having specific models (for each gender or rage). Obviously for this we require, first to be able to detect the faces and to implement accurate age, gender or race classification systems.

In this general context, we have developed a face analysis system, which can be used in the construction of human-computer interaction applications. The proposed face analysis system can detect faces, detect eyes, and classify the faces in dynamic environments with high robustness and speed. Our face detector, eye detector and gender classifier are built using a unified learning framework based on nested cascades of boosted classifiers. In the next subsections will be described this learning framework, the employed low-level features, and the resulting detectors and classifiers.

2.1 Learning Framework

Key concepts used in the learning framework are boosting [32], nested cascade classifiers [30], and bootstrap training [33]. A detailed description of this framework is given in [17].

Boosting is employed for finding (i) highly accurate hypotheses (classification rules) by combining several weak hypotheses (classifiers), each one having a moderate accuracy, and (ii) self-rated confidence values that estimate the reliability of each prediction (classification). Cascade classification uses several layers (stages) of classifiers of increasing complexity (each layer discards non-object patterns) for obtaining an optimal system in terms of classification accuracy and processing speed [29]. This is possible because of two reasons: (i) there is an important difference in the a priori probability of occurrence of the classes, i.e. there are much more non-object than object patterns, and (ii) most of the non-objects patterns are quite different from the object patterns, therefore they can be easily discarded by the different layers. Nested cascade classification allows obtaining higher classification accuracy by the integration of the different cascade layers [30]. Other aspects employed in the proposed framework for obtaining high-performance classification systems are: using the bootstrap procedure [33] to correctly define the classification boundary, LUTs (Look-Up Tables) for a fast evaluation of the weak classifiers, simple rectangular Haar-like features that can be evaluated very fast using the integral image [29], and LBP features [31] that are invariant against changing illumination.

2.2 Boosted Nested Cascade

A nested cascade of boosted classifiers is composed by several integrated (nested) layers, each one containing a boosted classifier. The whole cascade works as a single classifier that integrates the classifiers of every layer. A nested cascade, composed of *M* layers, is defined as the union of *M* boosted classifiers H_C^k each one defined by:

$$H_C^k(x) = H_C^{k-1}(x) + \sum_{t=1}^{T_k} h_t^k(x) - b_k \qquad (1)$$

with $H_C^k(x) = 0$ and h_t^k the weak classifiers, T_k the number of weak classifiers in layer k, and b_k a threshold value. It should be noted that a given classifier corresponds to the nesting (combination) of the previous classifiers. The output of H_C^k is a real value that corresponds to the confidence of the classifier and its computation makes use of the already evaluated confidence value of the previous layer of the cascade.

Each weak classifier is applied over one feature computed in every pattern to be processed. The weak classifiers are designed after the domain-partitioning weak hypotheses paradigm [32]. Under this paradigm the weak classifiers make their predictions based on a partitioning of a feature domain F. A weak classifier h will have an output for each partition block, F_j, of its associated feature f: $h(f(x)) = c_j \ni f(x) \in F_j$. Thus, the weak classifiers prediction depends only on which block F_j a given sample (instance) falls into. For each classifier, the value associated to each partition block (c_j), i.e. its output, is calculated for minimizing a bound of the training error and at the same time a bound on an exponential loss function of the margin [32]. This value is given by:

$$c_j = \frac{1}{2}\ln\left(\frac{W_{+1}^j + \varepsilon}{W_{-1}^j + \varepsilon}\right) \qquad (2)$$

$$W_l^j = \Pr[f(x) \in F_j \wedge y_i = l], l = \pm 1 \qquad (3)$$

and ε a regularization parameter.

2.3 Features

We use rectangular and modified LBP features for our classifiers. Rectangular features resemble Haar wavelets and can be evaluated very quickly, independently of their size and position, using the integral image [29]. They correspond to the difference between sums of pixels values in rectangular image regions. The output value defines a domain that is partitioned using intervals (or bins) of equal size [30]. The LBP (Local Binary Pattern) corresponds to an illumination invariant descriptor of the local structure in a given image neighborhood. We use their modified version [31] (mLPB = modified LBP), which overcomes some problems of the original LBP. The mLBP is computed as follows: for a given window of 3x3 pixels, the average of the pixels in the window is calculated. Then, each of the pixel values is compared against the obtained average. From these comparisons, 9 bits are generated, with 0 indicating that a pixel is smaller than the average, and 1 otherwise. After that, the concatenated 9 bits corresponds to the mLBP feature.

2.4 Detection and Classification Systems

Our face detection system works as follows: First, to detect faces at different scales, a multiresolution analysis of the images is performed, by downscaling the input image by a factor of – e.g. 1.2 --. This scaling is performed until images of about 24x24 pixels are obtained. Afterwards, windows of 24x24 pixels are extracted for each of the scaled versions of the input image. The extracted windows could then be pre-processed to obtain invariance against changing illumination, but thanks to the used of illumination invariant features we do not perform any kind of preprocessing. Afterwards, the windows are analyzed by a nested cascade classifier, built using the framework described in the former subsection. Finally, the windows classified as faces are fused (normally a face will be detected at different scales and positions) to obtain the size and position of the final detections. This fusion is described in [12].

The eye detector works in the same way as the face detector does, the only difference is that the search is not done within the whole image, but only within the face. As the face detector, the eye detector woks on 24x24 windows, therefore it can be used only on faces of 50x50 pixels or larger. The gender classifier was built using the learning framework as the eye and face detectors. The gender classifier works on windows of 24x24 pixels and when the eye positions are available it uses them for aligning the faces. In [5] we give a detailed description of the gender classifier.

For testing purposes we employed four databases BIOID [42], FERET [36], CMU-MIT [37], and UCHFACE (available in [2]). No single image from these databases was used for the training of our systems. Selected examples of our face detection, at work in the FERET, BIOID, UCHFACE and MIT-CMU databases, are shown in figure 1. This figure also shows eyes detection and gender classification.

The face detector was evaluated using two types of databases: (a) BIOID and FERET, which contain one face per image, and (b) CMU-MIT and UCHFACE, which contain none, one or more faces per image. Table 1 shows results of our method for the FERET, BIOID and UCHILE databases as well as the results of [31] for the BIOID database. In the BIOID database, which contains faces with variable expressions and cluttered backgrounds, we obtain a high accuracy, a 94.1% detection rate with zero false positives (in 1,521 images), while on the FERET database, which contains faces with neutral expression and homogeneous background, we obtain a very high accuracy, a 99.5% detection rate with 1 false positive (in 1,016 images). These results were obtained without considering that there is only one face per image. In the UCHFACE database (343 images), which contains faces with variable expressions and cluttered backgrounds, we consider that the obtained results are rather good (e.g. 88.0% with 3 false positives, 98.5% with 17 false positives).

Table 1: Comparative evaluation (DR: Detection Rate vs. False Positives) of the face detector on the BioID Database (1,521 images).

False Positives	**DB**	**0**	**1**	**2**	**3**	**6**	**15**	**17**	**25**
Out Method	UCHILE			87.8	88.0	94.8		98.5	
Out Method	FERET	98.7	99.5		99.7				
Out Method	BIOID	94.1	95.1	96.5		96.9	97.6		98.1
Fröba and Ernst 2004 [31]	BIOID		~50			~65	~84		~98

Table 2: Comparative evaluation (DR: Detection Rate vs. False Positives) of the face detector on the CMU-MIT database (130 images, 507 faces). Notice that in (Fröba et al. 2004) a subset of 483 (out of 507) faces is considered. This subset is called CMU 125 test set.

False Positives	**0**	**3**	**5**	**6**	**10**	**13**	**14**	**19**	**25**	**29**	**31**	**57**	**65**
Our Method		77.3	83.2				86.6	88	89.9		90.1		92.1
Fröba et al. 2004 [31]	~66		~87						~90				
Wu et al. 2004 [30]		**89**			90.1	**90.7**						**94.5**	
Viola and Jones 2001 [29]					76.1						88.4		92
Rowley et al.1998					83.2						86		
Schneiderman 2004				**89.7**				**93.1**		**94.4**			
Li. et al. 2002					83.6						90.2		
Delakis and Garcia 2004	**88.8**				**90.5**						91.5		92.3

Table 2 shows comparative results with state of the art methods for the CMU-MIT database. In the CMU-MIT database we also obtain good results (e.g. 83.2% with 5 false positives and 88% with 19 false positives). If we compare our results with state of the art methodologies in terms of DR and FP, we obtain better results than [29][37], slightly better results than [35], slightly worse results than [34] (but our system is about 8 times faster), and worse results than [30] and [38], mainly because of the size of the training database. For example in [30] 20,000 training faces are employed. while our training database consists of just 5,000 face images. Notice that our classifier is among the fastest ones. The ones that have a comparable processing time are [29], [31], [30], and [35].

The gender classifier performance was evaluated in two cases: when the eyes were manually annotated and when the eyes were automatically detected. Table 3 shows results of this evaluation for the UCHFACE, FERET and BIOID databases. It should be noticed that its behavior is very robust to changes in the eyes positions that are used for the face alignment, and that in two of the databases best results are obtained when the eye detector is used.

An evaluation of the performance of our eyes detector can be found in [17].

Table 3: Gender classification results: Percentage of correct classification when eyes are annotated or detected.

Database	Annotated eyes	Detected eyes
UCHFACE	81.23 %	80.12%
FERET	85.56 %	85.89%
BIOID	80.91 %	81.46%

3 Bender: A Personal Robot that Provides Entertainment, Companion and a Communication Interface

Personal robots are becoming of increasing interest in the robotics community. A personal robot is a subclass of a mobile service robot designed to interact with humans and to behave as a partner. It is usually expected that the personal robot's morphology and dimensions allow him to adequately operate in human environments. Among other abilities personal robots should be able to: (i) move in human environments, (ii) interact with humans using human-like codes (speech, face and hand gestures), (iii) manipulate objects, (iv) determine the identity of the human user (e.g. "owner 1", "owner 2", "unknown user", "Peter") and its mood (e.g. happy, sad, excited) for personalizing its services, (v) store and reproduce digital multimedia material (images, videos, music, digitized books), and (vi) connect humans with data or telephone networks. In addition, (vii) they should be empathic (humans should like them), and (viii) their use should be natural and should not require any technical or computational knowledge. In addition, thanks to the fact that personal robots are empathic and designed to be used by the non-expert, their application as Web interfaces can expand the use of Web and Internet technologies to humans that normally do not use computers (e.g. elderly people), and they can expand also the use of these technologies to situations where currently they are not used. Thus, it is expected that personal robots can provide entertainment, companion and communication interfaces.

Having these ideas in mind we have designed Bender, a service robot that can provide all these services (see figure 2). As a communication interface Bender provides multimedial communication (audio, video and text), natural use (speech, face interaction, hand-gesture interaction, touch screen on the robot's chest), selfpersonalization (face recognition + facial expression, gender and age recognition), empathy (emotions expression) and mobility [27]. Many of these features are achieved using a face analysis subsystem, which was built using the framework described in the former section.

Bender participated successfully in the RoboCup 2007 @Home competition, obtaining the *RoboCup @Home 2007 Innovation Award* for its human – robot interaction and object recognition abilities.

Figure 1. Selected examples of our face detection, eyes detection and gender classification systems at work on the MIT-CMU (a), FERET (b), UCHFACE (c) and BIOID (d) databases. In (e) and (f) are shown examples of detections for the system running in an AIBO robot. The system detect faces and performs gender classification. When the resolution of the faces is larger than 50x50 pixels it detects also the eyes.

Figure 2. The personal robot Bender.

4 Conclusions

In this article we have presented our research group in computer vision at the Universidad de Chile, as an example of how difficulties derived from carrying out research in a developing country can be overcome through international cooperation, focusing on global research projects with low requirements of expensive equipments, formation of the qualified researchers, synergistic research with other fields, and some necessary basic support from government research agencies. We have show concrete examples of how, we have been able to contribute in research areas such as face analysis for human-computer interfaces, robust object recognition, and robot vision, at an international level, been a small research group.

Acknowledgements

This research was partially funded by Millennium Nucleus Center for Web Research, Grant P04-067-F, Chile. Special thanks to all students from my research groups that have collaborated in this research, specially to Rodrigo Verschae, Patricio Loncomilla, Mauricio Correa, Pablo Guerrero, Paul Vallejos and Rodrigo Palma.

References

[1] UChile Official Website: http://www.robocup.cl/

[2] Computational Vision Lab (Universidad de Chile) Website: http://vision.die.uchile.cl/

[3] Ruiz-del-Solar, J., Loncomilla, P., Vallejos, P. (2007). An automated refereeing and analysis tool for the Four-Legged League. *Lecture Notes in Computer Science* 4434 (*RoboCup 2006*), pp. 206-218..

[4] Loncomilla, P., and Ruiz-del-Solar, J. (2006). A fast probabilistic model for hypothesis rejection in SIFT-based Object Recognition. *Lecture Notes in Computer Science* 4225 (*CIARP 2006*), pp. 696 - 705.

[5] Verschae, R., Ruiz-del-Solar, J., Correa, M. (2006). Gender Classification of Faces using Adaboost. *Lecture Notes in Computer Science* (*CIARP 2006*) 4225, pp. 68 - 78.

[6] Loncomilla, P., Ruiz-del-Solar, J. (2006). Gaze Direction Determination of Opponents and Teammates in Robot Soccer. *Lecture Notes in Computer Science* 4020 (*RoboCup 2005*), pp. 230 – 242.
[7] Loncomilla, P., Ruiz-del-Solar, J. (2005). Improving SIFT-based Object Recognition for Robot Applications. *Lecture Notes in Computer Science* 3617 (*ICIAP 2005*), pp. 1084 - 1092.
[8] Ruiz-del-Solar, J., Verschae, R., Kottow, D. (2005). Skin Detection in Videos in the Spatial-Range domain. *Lecture Notes in Computer Science* 3617 (*ICIAP 2005*), pp. 961-969.
[9] Ruiz-del-Solar, J., and Navarrete, P. (2005). Eigenspace-based Face Recognition: A comparative study of different approaches. *IEEE Trans. on Systems, Man, and Cybernetics C* (Special Issue on Biometric Systems), Vol. 35, No. 3, pp. 315 – 325, Aug. 2005.
[10] Ruiz-del-Solar, J., and Vallejos, P. (2005). Motion Detection and Tracking for an AIBO Robot using Camera Motion Compensation and Kalman Filtering. *Lecture Notes in Computer Science* 3276 (*RoboCup 2004*), pp. 619-627.
[11] Zagal, J., Ruiz-del-Solar, J., Guerrero, P. and Palma R. (2004). Evolving Visual Object Recognition for Legged Robots. *Lecture Notes in Computer Science* 3020 (*RoboCup 2003*), pp. 181-191.
[12] Verschae, R., and Ruiz-del-Solar, J. (2003). A Hybrid Face Detector based on an Asymmetrical Adaboost Cascade Detector and a Wavelet-Bayesian-Detector. *Lecture Notes in Computer Science* 2686 (*IWANN 2003*), pp. 742-749.
[13] Navarrete, P., and Ruiz-del-Solar, J. (2002). Analysis and Comparison of Eigenspace-based Face Recognition Approaches. *Int. Journal of Pattern Recognition and Artificial Intelligenc*e, Vol. 16, No. 7, pp. 817-830.
[14] Palma-Amestoy, R., Guerrero, P., Vallejos, P., and Ruiz-del-Solar, J. (2006). Context - dependent color segmentation for Aibo robots, *3rd IEEE Latin American Robotics Symposium – LARS 2006*, Oct. 26 - 27, Santiago, Chile (CD Proceedings).
[15] Ruiz-del-Solar, J., Loncomilla, P., Vallejos, P. (2006). An automated refereeing and analysis tool for robot soccer, *CVBASE '06 - Workshop on Computer Vision Based Analysis in Sport Environments* (*ECCV 2006 Workshop*), Graz, Austria, May 12, 2006.
[16] Verschae, R., Ruiz-del-Solar, J., Köppen, M., García, R. V. (2005). Improvement of a Face Detection System by Evolutionary Multi-Objective Optimization, *5th Int. Conf. on Hybrid Intelligent Systems - HIS 2005*, pp. 361-366, Nov. 6-9, 2005, Rio de Janeiro, Brazil.
[17] Verschae, R., Ruiz-del-Solar, J., Correa, M. (2007). A Unified Learning Framework for object Detection and Classification using Nested Cascades of Boosted Classifiers. *Machine Vision and Applications* (in press).
[18] Ruiz-del-Solar, J. and Verschae, R. (2004). Robust Skin Segmentation using Neighborhood Information. *Int. Conf. on Image Processing - ICIP 2004*, pp.207 – 210, Oct. 24 – 27, Singapore.
[19] Ruiz-del-Solar, J. and Verschae, R. (2004). Skin Detection using Neighborhood Information. *6th Int. Conf. on Face and Gesture Recognition – FG 2004*, pp. 463 – 468, Seoul, Korea, May 2004.
[20] Ruiz-del-Solar, J., Shats, A. and Verschae, R. (2003). Real-Time Tracking of Multiple Persons. *Proc. of the 12th Int. Conf. on Image Analysis and Processing – ICIAP 2003*, pp. 109 – 113, September, Mantova, Italy.
[21] Guerrero, P., Ruiz-del-Solar, J., and Palma-Amestoy, R. (2007). Spatiotemporal Context in Robot Vision: Detection of Static Objects in the RoboCup Four Legged League. *Proc. Int. Workshop on Robot Vision*, in *2nd Int. Conf. on Computer Vision Theory and Appl. – VISAPP 2007*, pp. 136 – 148, March 8 – 11 2007, Barcelona, Spain..
[22] Ruiz-del-Solar, J., Verschae, R., Vallejos, P., and Correa, M. (2007). Face Analysis for Human Computer Interaction Applications, *2nd Int. Conf. on Computer Vision Theory and*

Appl. – VISAPP 2007, Special Sessions, pp. 23 – 30, March 8 – 11 2007, Barcelona, Spain..

[23] Arenas, M., Ruiz-del-Solar, J., Verschae, R. (2007). Detection of Aibo and Humanoid Robots using Cascades of Boosted Classifiers. *RoboCup Symposium 2007* (in press).

[24] Guerrero, P., Ruiz-del-Solar, J., Fredes, J., and Palma-Amestoy, R. (2007). Automatic On-Line Color Segmentation using Classes-Relative color Spaces. *RoboCup Symposium 2007* (in press).

[25] Guerrero, P., and Ruiz-del-Solar, J. (2007). Improving robot self-localization using landmarks' poses tracking and odometry error compensation. *RoboCup Symposium 2007* (in press).

[26] Loncomilla, P., and Ruiz-del-Solar, J. (2007). Robust Object Recognition using Wide Baseline Matching for RoboCup Applications. *RoboCup Symposium 2007* (in press).

[27] Ruiz-del-Solar, J. (2007). Personal Robots as Ubiquitous-Multimedial-Mobile Web Interfaces, *LA-WEB 2007*, Santiago, Chile (in press).

[28] Ruiz-del-Solar, J., Loncomilla, P., and Devia, C. (2007). A New Approach for Fingerprint Verification based on Wide Baseline Matching using Local Interest Points and Descriptors. *Proc. PSIVT 2007* (in press).

[29] P. Viola and M. Jones, "Fast and robust classification using asymmetric Adaboost and a detector cascade", *Advances in Neural Inform. Processing System* 14, MIT Press, 2002.

[30] B. Wu, H. Ai, C. Huang, and S. Lao, "Fast rotation invariant multi-view face detection based on real Adaboost", *6th Int. Conf. on Face and Gesture Recognition - FG 2004*, pp. 79–84, Seoul, Korea, May 2004.

[31] B. Fröba and A. Ernst, "Face detection with the modified census transform", *6th Int. Conf. on Face and Gesture Recognition - FG 2004*, pp. 91–96, Seoul, Korea, May 2004.

[32] Schapire, R.E. and Singer, Y., 1999. Improved Boosting Algorithms using Confidence-rated Predictions, *Machine Learning*, 37(3), pp. 297-336.

[33] Sung, K. and Poggio, T., 1998. "Example-Based Learning for Viewed-Based Human Face Deteccion", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.20, N. 1, pp. 39-51.

[34] Delakis, M. and Garcia, C., 2004. "Convolutional face finder: A neural architecture for fast and robust face detection", *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(11), pp. 1408 – 1423.

[35] Li, S.Z., Zhu, L., Zhang, Z.Q., Blake, A., Zhang, H.J. and Shum, H., 2002. "Statistical Learning of Multi-view Face Detection", *7th European Conf. on Computer Vision – ECCV 2002*, (*Lecture Notes in Computer Science*; Vol. 2353), pp. 67 – 81.

[36] Phillips, P. J., Wechsler, H., Huang, J.and Rauss, P., 1998. "The FERET database and evaluation procedure for face recognition algorithms," *Image and Vision Computing J.*, Vol. 16, no. 5, pp. 295-306.

[37] Rowley, H., Baluja, S. and Kanade, T., 1998. "Neural Network-Based Detection", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol.20, No. 1, pp. 23-28.

[38] Schneiderman, H., 2004. Feature-Centric Evaluation for Efficient Cascade Object Detection, *IEEE Conference of Computer Vision and Pattern Recognition – CVPR 2004*, pp. 29 - 36.

[39] ALFA Program Official Website. Available electronically on August 2007 in: http://ec.europa.eu/europeaid/projects/alfa/index_en.htm

[40] Fraunhofer-Institute for Production Systems and Design Technology Official Website. Available electronically on August 2007 in: http://www.ipk.fraunhofer.de/en/

[41] Sicherung der Qualität von digitalen Gesichtbildern in *Futur* Magazin, Franunhofer-Institute IPK Berlin, Germany, No. 2, 2006.

[42] BioID, 2005.Face Database. Available on August 2005 in: http://www.bioid.com/downloads/facedb/index.php