Horopter based Dynamic Background Segmentation applied to an Interactive Mobile Robot

José Prado, Luis Santos, Jorge Dias

***Abstract*— Interactive mobile robots require object/subject detection in very visually complex environments. In the field of computer vision, specially when applied to robotics, several approaches like face detection, face recognition and pedestrian detection often have to deal with issues associated to bad illumination and strong featured background. These issues imply lack of performance because human detection algorithms will frequently process the whole image searching for features. Also, background segmentation approaches are commonly used to solve this problem on static camera surveillance. However all these approaches are unable to effectively deal with the constant background changes that certainly happen when the camera sensor is installed on a mobile robot. Hence, in this work we propose a *Horopter based Dynamic Background Segmentation* solution to this problem. Results show that our approach, significantly enhanced tracking, and consequently improved movement classification towards interaction.**

Fig. 1: Segway Robotic Mobility Platform (RMP) equipped with the robotic head. Configuration of the robot ready for human-interaction, gaze tracking and pursuit.

I. Introduction

Human-robot interaction approaches like face detection, face recognition [12], [16], [17], [13], pedestrian detection [11], [10] are widely known in robotics field; however frequently they lead to performance problems. Additionally, false positive and false negative problems are commonly associated to bad illumination. Interactive robots mostly interacts with the closest subject, another common assumption it that the robot will never deal with two subjects at the same time. Horopter is the optical phenomenon in stereo vision that happens in a form of a 3D curve on the Cartesian space where the disparity is zero (details in section II-C). Taking this into account, we propose a Horopter based Dynamic Background Segmentation (DBS) in order to reduce the searching space to a *zone of interaction*[1] [2]

To define *Zone-of-Interaction* we introduce the *horopter* definition. As it will be explained in section II-C, given three points (whose coordinates correspond to two cameras and the desired focus point), it is possible to draw a circumference containing all of them. The inside area of that circumference is the interaction zone. This means that only objects inside that area are possible of being detected, and thus to interact with the robot. Our approach is based on the *Geometric Horopter* and in order to calculate the horopter, first it is necessary to have the stereo *disparity map.*

Although we like to encourage the use of our approach to social robots, we prefer to call our robot an interactive-robot, since social robots might have several social aspects that our interactive-robot does not have (like facial expressions, arms and voice).

In this work, section II-A describes the calibration method for the stereo camera system, using the homography concept. Briefly, we can define homography [6] as a geometrical method which allows a linear transformation (using the homographic matrix) of coordinates between two planes. The following section II-C will show how the horopter calculation proceeds. Further in section II-D we give an example of how face and hand recognition frequently used on gesture recognition algorithms could have better results with our approach. In section II-E we explain how we did implement our robotic head tracker in order to have better interaction with humans. Finally in section III-A, as a study case, we implemented this technique to improve the results of a gesture recognition algorithm based on Laban movement analysis proposed in [14].

Featured base face detection: Recently Bau-Cheng Shen and Chu-Song Chen proposed a method to retrieve similar face images from large face databases. The proposed method extracts a set of Haar-like features, and integrates these features with supervised manifold learning. Haar-like features are intensity-based features. The values of various Haar-like features comprise the *rectangle feature vector* (RFV) (detailed on [16]), to describe faces. Compared with several popular unsupervised dimension reduction methods, RFV is more effective in retrieving similar faces. To further improve the performance, [16] combine RFV and a supervised manifold learning method and obtain satisfactory retrieval results.

Skin color hand detection: According to [8], skin color can provide a useful and robust cue for human-related image analysis, such as face detection, hand detection and tracking,

[1] *zone of interaction* is the region inside the horopter 3D space (see theoretical horopter definition on section II-C)

[2] The authors gratefully acknowledge support from EC-contract number BACS FP6-IST-027140, the contribution of the Institute of Systems and Robotics at Coimbra University and reviewers' comments.

people retrieval in databases and Internet, etc. The major problem of such kinds of skin color detection algorithms is that it is time consuming and hence cannot be applied to a real time system.

A. Related Work

Robotics has already acknowledged the evidence that human movements could be an important cue for Human-Robot Interaction. Sato [15], while defining the requirements for 'human symbiosis robotics' state that those robots should be able to use non-verbal media to communicate with humans and exchange information. As input modalities on a higher abstraction level they define channels on language, gesture and unconscious behavior. This skill could enable the robot to actively perceive human behavior, whether conscious and unconscious. Human intention could be understood, simply by observation, allowing the system to achieve a certain level of friendliness, hospitality and reliance. By using a reference image, a video coding approach has previously been developed in the context of road surveillance [18]. Moreover it was shown how the image reference was built during initialization phase.

The classical background subtraction technique was used to perform the segmentation of mobile objects. Instead of updating the remote reference with a specific period, [18] presented a technique to update the remote background image by pieces. The updating of the remote reference is triggered when some specific conditions are met, depending on the amount of moving areas. In [9] an integrated system for smart encoding in video surveillance was presented. Their system aims at defining an optimized codestream organization directly based on the semantic content of the video surveillance analysis module. The proposed system produces a fully compliant motion stream that contains regions of interest (typically mobile objects) data in a separate layer than regions of less interest (e.g. static background). It can also be used in applications requiring selective scrambling of regions of interest as well as for any other application dealing with regions of interest.

B. Contextualization

There is a pre-requisite on the field of human-robot interaction, this would be the need for the robot to recognize the person with whom it will interact. Usually it is done using a video sensing. Since the system is implemented in a mobile platform, to separate the person from the background demands more complex processing, due to dynamic characteristics of the background. This means that an approach based in static background, as in [18] and [9], is not possible. The challenge was thus to have a robust real time solution for dynamic background segmentation on mobile robotics.

Our approach is then based on the *Geometric Horopter* as will be shown in section II-C. Our interactive robot shown in Fig. 1[3] will consider visible objects only if they are inside the *zone of interaction* region (projected on 2D space of camera image plane). The expected result is seen in figure 2a, where the acquired image from the robot's perspective can be seen. However, situations occur where interference exists. In figure 2b we see a multi-person (noisy) scenario, where the presence of a second subject would interfere on the analysis of several algorithms [12], [16], [17], [13], [11], [10]. Our approach, DBSH (Dynamic Background Segmentation based on Horopter), will only detect subject inside the horopter zone, which is represented by a dashed line at the floor (figure 2b).

[3]The robotic head was developed in University of Coimbra Portugal with support from POP European project and Professor Dr Helder Araujo

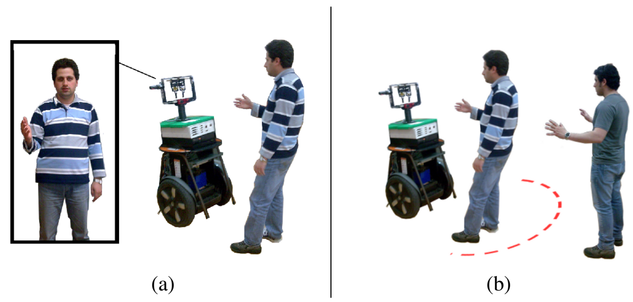


Fig. 2: a) Image acquisition from the robot's point of view b) Noisy scenario, another subject trying to interfere during the interaction

II. Our Approach

A. Camera Calibration

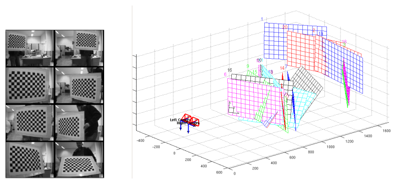
Camera calibration has been extensively studied, and standard techniques established. For this work, camera calibration was performed using the Camera Calibration Toolbox for Matlab [2]. The C implementation of this toolbox is included in the Intel Open Source Computer Vision Library [7].

The calibration uses images of a chessboard target in several positions and recovers the camera's intrinsic parameters, as well as the target positions relative to the camera, as shown in fig 3b. The calibration algorithm is based on Zhang's work in estimation of planar homographies for camera calibration [20], but the closed-form estimation of the internal parameters from the homographies is slightly different, since the orthogonality of vanishing points is explicitly used and the distortion coefficients are not estimated at the initialization phase. The calibration toolbox will also be used to recover camera extrinsic parameters and homographic matrix between the two cameras of the stereo system.

B. Disparity map and Depth map

Disparity maps represents the difference distance between points[4] of a pair of images; meanwhile *depth maps* represents the expected depth/distance that an area is considered to be away from the camera.

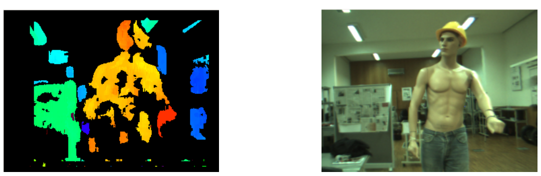
[4]these points can be either raw pixels or features depending on the approach



(a) (b)

Fig. 3: Stereo camera calibration with Bouguet Matlab toolbox
a) Images with chessboard target used for calibration b)Reconstructed target positions relative to the camera frame referential

We used an adaptation[5] of the Videre [19] libraries in order to get the depth map. Videre library first construct a disparity space image from stereo image pair, and then calculate temporary disparity maps using the SAD method [4]. Later stage of the algorithm will reduces both the blurred errors at depth discontinuities and the mismatched errors at half occluded areas. The final step is to use a median filter to interpolate the dense disparity map. Once one has calibrated the cameras and the *disparity map* calculated, it is trivial to get the *depth map* Fig.4a a).



(a) (b)

Fig. 4: a) Depth map (*'hot'* colors represent nearest areas, *'cold'* colors represent further ones b)Dominant eye raw image

C. Geometrical Horopter

1) Properties of ViethMuller Circle: The concept of *interaction zone* has been defined as dependent of a circle. That circle is called the Vieth-Muller Circle. The Vieth-Muller Circle defines the region where the disparity is equal to zero, while the disparity grows for inside with positive values and grows (shrink if considering the raw value) to out-side with negative values.

Pixels that present negative values for disparity, will be assigned zero value (black color pixels). The result is a segmented image where the pixels calculated to be inside the *Vieth-Muller* circle define the 'visible' objects within the circle (the interaction zone). The segmented image (right column of figure 5) results in a region of interest and this region will define the true input pixels for the *face/hand* detector. Consequently the robot will interact only with subjects inside *Vieth-Muller* circle, *i.e.* inside its current horopter.

Notice that we still have some noise at the segmented images, these noisy areas exists usually due to homogeneous areas in the original image. Homogeneous areas and also very similar neighbor features of the image can add noise to our depth map and consequently to the final horopter segmented image. Although we have this noise the result is still better for hand and face detection than if you have no segmentation.

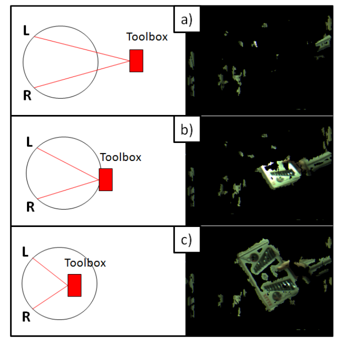


Fig. 5: a) The toolbox is yet outside the Vieth-Muller Circle; b) Toolbox starting to enter the horopter zone; c) The object is fully inside the Vieth-Muller circle, and thus, visible.

D. Subject Detection

Our system performs subject detection starting with face and hand detection. Further it combines these features using a body shape triangle representation as described on subsection III-A.1. Once the body-shape triangle is defined, it is possible to assign the triangle properties with the LMA expected variables according to table I.

Additionaly we combined face and hand detection algorithms with the horopter dynamic segmentation. We firstly do the dynamic background segmentation, hence it is only necessary to slide on the remaining pixels; this significantly increases the detection performance. Thus we have very fast (10 fps) results on the segmentation plus detection. The gesture recognition algorithm proposed in [14] assumes always the same default initial position for face and hands, later on the process it tracks the real position; this approach implies on performance lost during godfather[6] localization. Thus, in order to save start up time, our choice was to firstly detect the face and the hands position with the algorithms previously mentioned and give this as input to the gesture recognition algorithm.

The red oval on *fig. 9 b)* is an approximation of the search region. It is observable on the *right b)* image that there are areas with skin color on the wall and floor, so if the full image was passed to the hand algorithm hand false positives would certainly occurs. Furthermore similar errors could happen for the face algorithm if the background was strongly and randomly featured.

[5]Videre cameras usually do not vary so much the interocular distance from one model to another, that fact makes possible to have reasonable results even without calibration. In our case we adapted this library by allowing grabbing from other cameras. We also allowed external calibration data input; thus we used in this work two monocular cameras in a stereo system that allows vergence.

[6]godfather is defined on [14] as the person whom the robot is supposed to interact

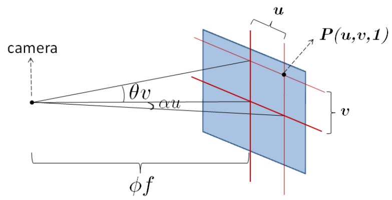


Fig. 6: αu= pan ,θv = tilt — Tracking angles to the robotic head

E. Tracking

If a subject remains inside horopter for some seconds, the robot will elect this subject to be it's godfather. Let's call godfather the human elected for interaction with the robot. Hence the robot locates his face and hands as explained in section II-D. As our robot is an interactive robot, we want it to track the godfather while he moves also. In order to have an intuitive interaction it is necessary that the man sees the robot facing him; or, in other words, the robotic head needs to move targeting at the center of the subject head.

In homogeneous coordinates consider an image point $P(x, y, z, 1)$, after normalization $P(u, v, 1)$; knowing focal length f from camera intrinsic calibration, d is the distance from the camera to the target object, and an empirically found multiplier λ. We have: $\phi = d * \lambda$. Due to the fact that u and v are initially in pixels while d and f are initially in centimeters, the multiplier λ is necessary.

Then, finally we have as it is visible on fig. 6 $tan\theta v = \frac{v}{\phi f}$ and $tan\theta u = \frac{u}{\phi f}$.

III. Applications of Dynamic Background Segmentation to Interactive Robotics

As mentioned on previous sections one of the principles we are focused in is *interaction*. The interaction scheme can be simplified and thus divided in two stages:*Whom* to interact with; *How* to interact. The whom question as been described throughout sections II-C to II-E. This section will give a general overview on the how.

A. Laban Movement Analysis

Rett J. in his work [14], investigated the possibility of using Laban Movement Analysis (LMA) to classify human movements. Laban Movement Analysis, is a descriptive language of dancing movements. It was developed by Rudolf Laban (1879 to 1958), considered by many a pioneer of European modern dance and theorist of movement education. There are some studies related to LMA, but this is particularly interesting, because an interactive robot was developed to serve as a demonstrator of the usability of this technique.

Literature is not in consensus about the number of LMA components. Most notably, the work of Norman Badler's group [3], [21], [22], [1], divides LMA into five components (Fig. 7) that are: *relationship, space, body, shape, effort*. Each of this latter four components deals with a specific aspect of movements. **Non kinematic** components: *Body* specifies which body parts are moving, their relation to the body center; *Space* deals directly with the trajectory executed by the body parts while performing a movement. Within the **Kinematic** ones there are: *Effort* which deals with the dynamic qualities of the movement, and the inner attitude towards the use of energy; *Shape* (emerging from *Body* and *Space*) is focused on the body itself. Then we have *Relantionship* that appears as the less explored component, and describes the interaction with oneself, others and the environment. Some literature only considers the first four mentioned components [5].

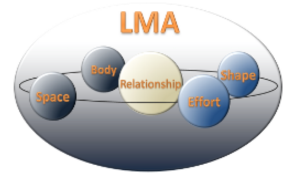
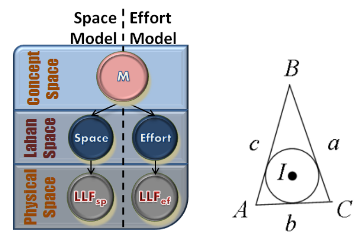


Fig. 7: The five LMA components



(a) (b)

Fig. 8: a) LMA Global Model b)The triangle formed by the hand-head-hand positions is used to express Shape

1) Laban Movement Analysis within robot-human interaction context: Each the mentioned components deals with a specific aspect of expressive movements, which are closely related to physical entities. The initial hypotheses of correspondences between some LMA parameters and physical entities are expressed as shown in I.

For the description of the Space component a feature based in displacement angles was chosen. This physical measurable entity represents the Space component of LMA very well and the process of computation is simple. When using a low cardinality we can expect a good performance of the Bayesian method for learning and classification.

Displacement angles, which also have been used by [21]

LMA parameter	Physical entities
Space	Displacement angle
Effort.Time(sudden)	High acceleration, High velocity
Effort.Time(sustained)	Low acceleration, Low velocity
Effort.Space(direct)	Small curvature, Small angular velocity
Effort.Space(indirect)	High curvature, High angular velocity
Effort.Weight(strong)	Muscle tension, Medium acceleration
Effort.Weight(light)	Muscle relaxed
Effort.Flow(free)	High curvature, High angular velocity
Effort.Flow(bound)	Low acceleration, Low velocity
SpatialShaping	Displacement angle
ShapeFlow	Position and shape of hand-head-hand triangle

TABLE I: Initial hypotheses of correspondences between LMA parameters and physical entities can be calculated easily from two subsequent positions. They describe the trace of a curve quite well and are independent from the absolute positions. As the position data is projected to planes, each plane produces a sequence of displacement angles with a certain sampling rate and discretization. For the Effort component of LMA, the assumption of an high acceleration when Time.sudden occurs seems to be a logical choice. The high velocity might follow as a consequence of the high acceleration. The inverse situation is assumed during Time-sustained when low acceleration and velocity is assumed. Interpreting Space-direct as reaching towards a target we can assume a straight trajectory of the hand. This suggests to take the curvature into consideration as a measure of 'directness'. The mathematical definition of curvature though, requires a parametrized curve which is independent of time t. The curvature k is approximated by calculating the change of displacement angles (angular change; angular velocity). The Weight quality is related with muscle tension but the attempt to give a measure for this can not be achieved within the scope of paradigm three. The only relation we draw is that highly tensed muscles can not exert high acceleration. The Flow quality is interpreted as the attempt of consciously following a planned trajectory or not. We assume that a Flow-bound movement will have a low acceleration and velocity, while in Flow-free a high curvature is expected. As it can be seen the mapping is not perfectly one-to-one and for Weight-light no (feasible) evidence is given at all.

While these two components (Space and Effort) that have just been described have already been implemented in [14], some space is left for improvement in the remaining components.

Regarding the body component, it deals with body itself and what body parts are moving related to the body center. As we consider only upper body expressive movements, the sternum is considered to be the body center. Body can use descriptors like spereading and skinking.

For the Shape component, face and hands can be combined in order to have a body shape and according to [14] the angle formed in the vertex of the head, can be used to recover a shape component. This will give the possibility to semantic descriptions like growing or shrinking. To strengthen our feature set, the perimeter of the triangle can also provide good information. The general geometrical concept behind this descriptors can be found in Fig.8b. Two measures are calculated one is the vertical position of the Incenter Iz of the triangle formed by the hands and the head. Upward displacements relative to the initial Incenter are indicating growing, while downward displacements indicate shrinking. Additionally, the total length of the triangle $l = a + b + c$ is used as an evidence for shrinking and growing.

B. Interaction

In [14], a Bayesian framework is used as support to the implementation of LMA. The Bayes net implementation is out of the scope of this work, however, Fig. 8a presents the global model for contextualization purposes. Since LMA is composed of four main components, Bayesian approach gives us the flexibility of component integration, i.e. each component can be modeled separately and integrated in a final global model. Also probabilistic approaches allow us to deal with uncertainty and incomplete data, which may also occur, in case tracking fails at some point. As input to the Bayesian network, features are provided as evidences. While movements are being performed, the tracking of body parts generate 2-D trajectories. The features (e.g. angle displacements, vectorial displacements, acceleration, etc.) emerge directly from these tracked trajectories.

Movement	Interpretation	Action
Circle	turn 360º	Rotation
Pointing	Acknowledgment	Perform Action
Wave Left	Step aside (left)	Move Left
Wave Right	Step aside (right)	Move right
Sagittal Wave	Come closer	Move Forward
Bye-bye	Ignore Gesture, Stop intercation	Switch system off

TABLE II: Movement and correspondent robot actions

A set of movements was learned, and a set of actions was assigned in response, i.e. the robot, through the probabilistic approach, estimated a determined movement through inference of the features, and consequently would react to its assumption. Table II shows the movements and the action responses.

As it can be seen, the actions of the robot are a direct consequence of the movement identification, and this identification relies on the robustness of the tracking algorithm.

As already previously stated, when using color tracking schemes, the tracker sometimes loses the target by means of generating false positives for body part identification. This is due to multi-colored backgrounds, which are very common within dynamic scenarios. Thus, by applying the geometric horopter technique to the system used in [14] was able to reduce the search area within the image. The perfect scenario occurs when a perfect bounding box around the human silhouette is generated, as it was theoretically represented on Fig. 2a. The algorithm slowed its tracking computational time, from deploying 15 frames/second to 10 frames/second, which is not considered critical, as 10 frames is still a good rate. This happened because the old version used one camera only, and after the application of this method, most processing time is dedicated to the computation of the depth image. However tracking results increased dramatically, by reducing the tracking false positives in 81%. To strengthen our tracking rate, geometric constraints were also applied.

As a direct consequence of the tracking enhancement, the movement classification also improved. Past experiments, with the old tracking method, often showed tracking deficiencies. This fact would lead to the necessity of performance repetition by the subject, so the movement could be recognized. The tracker, in complex environments would return false positives 43% of the time, hence, the movement would be misclassified. The approach based on the geometric horopter, improved the classification rate to 83.4% (from a past rate of 63.5%). These are significative results, if one understands that robot-human interaction should occur

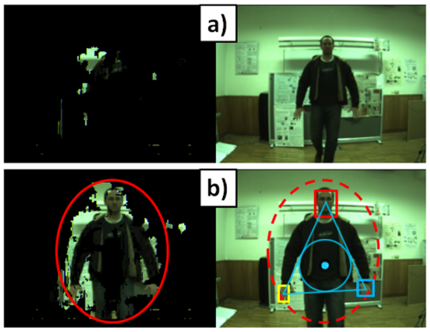


Fig. 9: a) Subject entering in horopter, consequently entering in the field of view of the robot. b) subject is inside the horopter and thus have his face and hands localized.

	Tracker loss	Movement C.R.	Real C.R.
Full image processing	43.0%	63.5%	36,2%
ROI processing	8.6%	83.4%	76,2%

TABLE III: Results for human-robot interaction.

smoothly. Problems like tracker loss and misclassification should be avoided, as this often results in the lack of interest from the human side.

Results are summarized in TableIII. Attention is called to the last column, where "real" classification rate (C.R.) present the results in a different perspective, so the reader understands the true enhancement of our approach. Considering that a movement is only classified when the tracker is not lost, i.e. tracker loss represents misclassification, the real classification rate for the old method returned positive identification of 36.2% for all trials. However, applying the new method (ROI stands for the Region Of Interest corresponding to the elipsoid around the subject's silhouette), results improved to 76.2%. This new perspective of looking at the results shows that our method effectively improved our classification rate in 110.5%.

Regarding movement classification alone, we can consider that LMA is a valid approach for movement classification. If body parts are correctly tracked, 83.4% of positive classifications is a good result, having in mind that not all components are yet implemented.

IV. Conclusion

Dynamic background segmentation is a good strategy to reduce the false positives of several algorithms that are based rather on pixel color or features. By reducing the scope of the searching image to an *zone of interaction* area, the applications of the DBS we proposed here are wide open on the field of Social Robots. In all the cases (haar like features face detection, skin color hand detection, gesture recognition with LMA), our DBS approach shown to improve the performance **and** the results.

It is known that gesture recognition and Laban Movement Analysis can provide us a good passive interaction. We concluded here that *horopter based dynamic background segmentation* can improve the performance, effectiveness and interactivity of the system in a mobile platform. Even thought, the concept of interactivity goes beyond reactive and passive movements. Taking this into account, we want to explore the area of learning and active behavior decision on robotics. Likewise, future work will certainly lead us to model a Bayesian network of behavior and stimulus, and to train this network using the approach for interaction presented on this paper.

References

[1] D. Bouchard and N. Badler. Semantic segmentation of motion capture using laban movement analysis. *In Proc. of Intelligent Virtual Agents*, 1:37–44, 2007.

[2] Jean-Yves Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calibdoc/index.html, 2006.

[3] Diane Chi, Monica Costa, Liwei Zhao, and Normal Badler. The emote model for effort and shape. *SIGGRAPH Computer Graphics Proceedings, Annual Conference Series*, 00:173–182, 2000.

[4] O. Faugeras, B. Hotz, H. Mathieu, T. Vie ville, Z. Zhang, P. Fuaand E. The ron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C Proy. Real time correlation-based stereo algorithm, implementations and application. Technical report, INRIA, Technical Report 2013, 1993.

[5] Afra Foroud and Ian Q. Whishaw. Changes in the kinematic structure and non-kinematic features of movements during skilled reaching after stroke: A laban movement analysis in two case studies. *Journal of Neuroscience Methods*, 158:137–149, 2006.

[6] R.I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[7] Intel. Intel open source computer vision library. http://www.intel.com/technology/computing/opencv/index.htm, 2006.

[8] Tarek M Mahmoud. A new fast skin color detection technique. In *World Academy of Science*, 2008.

[9] J. Meessen, C. Parisot, C. Lebarz, D. Nicholson, and J.F. Delaigle. Smart encoding for wireless video surveillance. In *In SPIE Proc. Image and Video Communications and Processing*, volume 1, 2005.

[10] L R Oliveira and Urbano Nunes. On integration of features and classifiers for robust vehicle detection. *IEEE Conference on Intelligent Transportation Systems (ITSC08)*, 2008.

[11] L R Oliveira and Urbano Nunes. On using cell broadband engine for object detection in its. *IEEE International Conference on Intelligent RObots Systems (IROS)*, pages 54–58, 2008.

[12] Minh Tri Pham and Tat Jen Cham. Fast training and selection of haar features using statistics in boosting-based face detection. *In Proc. 11th IEEE International Conference on Computer Vision (ICCV'07)*, 2007.

[13] Iain Matthews Ralph Gross and Simon Baker. Active appearance models with occlusion. *Image and Vision Computing*, 24:593–604, 2006.

[14] Joerg Rett. *Robot-Human Interface Using Laban Movement Analysis Inside a Bayesian Framework*. PhD thesis, University of Coimbra, 2009.

[15] Tomomasa Sato, Yoshifumi Nishida, and Hiroshi Mizoguchi. Symbiosis with human through behavior media robotics and autonomous systems. *Robotic Room*, 18:185–194, 1996.

[16] Bau-Cheng Shen, Chu-Song Chen, and Hui-Huang Hsu. Face image retrieval by using haar features. *Pattern Recognition ICPR*, pages 1–4, 2008.

[17] Barry-John Theobald, Iain Matthews, and Simon Baker. Evaluating error functions for robust active appearance models. *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, 1:149–154, 2006.

[18] Theodore Totozafiny, Olivier Patrouix, Franck Luthon, and Jean-Marc Coutellier. Dynamic background segmentation for remote reference image updating within motion detection jpeg2000. In *ICIP International Conference on Image Processing*, volume 1, 2008.

[19] Videre. Videre design. http://www.videredesign.com/.

[20] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. *In Proceedings of the Seventh International Conference on Computer Vision (ICCV' 99)*, 1:666–673, 1999.

[21] Liwei Zhao. Synthesis and acquisition of laban movement analysis qualitative parameters for communicative gestures. *PhD Thesis, University of Pennsylvania*, 2002.

[22] Liwei Zhao and Norman I Badler. Acquiring and validating motion qualities from live limb gestures. *Graphical Models*, 67:1–16, 2005.