Observing hand grasp type and contact points using hand distributed accelerometers and instrumented objects

Jorge Lobo, Pedro Trindade and Jorge Dias

***Abstract*— In order to study how humans grasp and manipulate objects diverse experimental setups have been used. Typically these hinder to some extent the natural movements and object interaction with the human hand. Our approach aims to minimise this interference by using minimalist sensing (i.e., only distributed tiny accelerometers on the hand), and instrumenting the manipulated object to have tactile data. Using MEMs tri-axial accelerometers on each fingertip and the palm, as well as on the object, relative angular pose can be determined by using gravity as a vertical reference. This can be used to identify the grasp type as well as the relative pose between the object and the hand. Preliminary results show the validity of the method, although the estimated relative angular pose is noisy, it is enough, together with the tactile data, to identify grasp types. Continuous observation as the overall pose changes for the same grasp type helps to overcome non-observability issues due to using gravity as a vertical reference.**

I. Introduction

The development of sophisticated robotic hands to work in manmade environments, capable of working with artefacts and objects designed to be manipulated by humans, requires a deep knowledge of how humans perform these tasks. A wide range of sensors and experimental setups can be used [1], however these typically also limit the free and natural hand manipulation of objects, or require complex setups only possible in the laboratory (multiple cameras, lasers, etc). Our approach aims to minimise this interference but at the same time have a device that is easy to use outside the lab. This will enable the study of a wider range of human manipulation tasks.

Distributed accelerometers can provide rich information about the orientation relative to the vertical gravity reference, as well as dynamic information about motion. Although the smartdust concept failed in a way to deliver the initially envisioned results, for sensing hand pose and motion, distributed accelerometers are an interesting solution [2] . The hand can be seen as a pice-wise rigid body with joint restricted movements, as well as some compliant parts. Minute sensors can be linked in a local bus and provide rich data on the pose and motion. Future implementations might even tap into the concept of energy harvesting, taking advantage of the hand kinetic energy to power the sensing or even the complete system.

Using MEMs tri-axial accelerometers on each fingertip and the palm, as well as on the object, relative angular pose can be determined by using gravity as a vertical reference. This can be used to identify the grasp type as well as the relative pose between the object and the hand. However some types of grasps are very dependent on the object geometry, and the force and contact surface is a more important cue than the finger orientation. By instrumenting the manipulated object with additional tactile sensors we can obtain these cues. The contact points provided by the tactile sensing on the object might also help to overcome non-observability issues due to hand motion and from using gravity as a vertical reference. The relative geometry of the fingers and the object is also a key parameter of the physical interaction that defines the grasp type, hence the importance of having a tri-axial accelerometer in the manipulated object.

In the scope of the HANDLE project [3] we have developed an experimental setup with diverse sensors shown in fig. 1. However the used Cyberglove, Polhemus magnetic tracker, and TekScan tactile sensors, whilst providing hand pose and tactile data, hinder to some extent the natural movement of the hand. In [4] our multi-sensor setup (fig. 1.) was used to record and analyse human in-hand manipulation. Although diverse and rich data is gathered, the glove and overlaid tactile sensing hinders the hand dexterity and natural movements, and the full sensor setup can only be used in well controlled lab conditions. To overcome these limitations, in our work we aim at having a clutter free and easily deployable system of identifying human natural manipulation of objects.

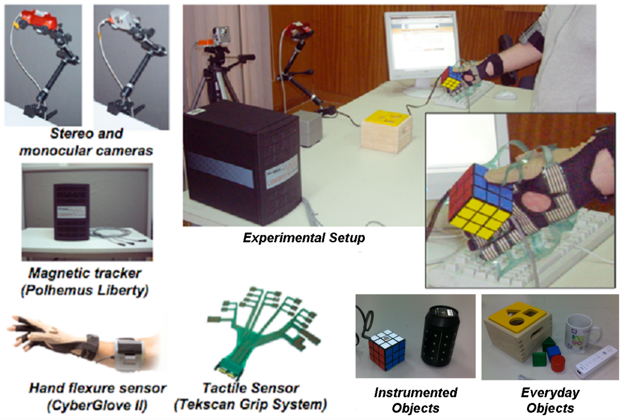


Fig. 1. Global overview of the HANDLE experimental area, data acquisition devices and objects available [3].

II. Related Work

Gesture recognition and man-machine interfaces (MMI) have pursued many ways of attaching sensors to the human hand, in [1] an extensive survey of glove-based systems their applications is presented.

An acceleration sensing glove was, to our knowledge, first presented in [2] as an input device for static gestures and as a pointing device for MMI. This short paper already envisions a future goal of Smart Dust on a finger, with self powered wireless accelerometers on the fingernails. Following the same idea, in [5] the *AcceleGlove* is presented as a whole-hand input device for virtual reality, although the focus is on MMI for mouse control and American Sign Language alphabet recognition.

In our previous work [6] hand distributed accelerometers from a commercial version of the AcceleGlove were used to identify static gestures, including the Portuguese Sign Language alphabet.

The research leading to these results has been partially supported by the HANDLE project, which has received funding from the European Community's 7th Framework Programme under grant agreement ICT 231640; and by the Institute of Systems and Robotics at Coimbra University.

J. Lobo, P. Trindade and J. Dias are with the Institute of Systems and Robotics DEEC - University of Coimbra - Polo II, 3030-290 Coimbra-Portugal {jlobo, pedrotrindade, jorge}@isr.uc.pt

The feature space consisted in the relative angular pose between each fingertip and the palm, and the roll and pitch of the palm. These were determined by using gravity as a vertical reference, and observing the same gesture in distinct poses. A simple nearest neighbour method identified the performed gesture against a library of gestures. This followed from our previous work on using gravity as a vertical reference for camera-IMU cross calibration [7] and in robot inertial aided vision [8].

While the focus of some of the above works is on gesture recognition and man-machine interfaces, we want to observe how humans grasp objects, and focus on grasp type classification. In hand motion research many taxonomies can be found. The most widely used grasp taxonomy is that of Cutkosky [9]. He focuses on all aspects of grasping, basing his taxonomy on that of Napier [10]. Van Nierop et al. introduced a hand-motion taxonomy in a two dimensional parameter space based on tasks that are evolutionary linked to the environment [11]. On the basis of a comparative literature research, the GRASP consortium developed a comprehensive human grasp taxonomy [12] [13]. A total of 33 different grasps were identified and arranged in an original taxonomy. The position of the thumb was introduced as an additional attribute, which can be either abducted or adducted. Depending on the need for precision, the taxonomy offers a second level of classification which includes only 17 grasp types. Following this taxonomy, in [14] a spatio-temporal modeling of grasping actions is presented.

III. Observing Relative Pose with Distributed Accelerometers

Assuming the hand is performing a steady grasp on the object, and distributed accelerometers on the hand and object (fig. 2), the vertical reference provided by gravity can be used to determine relative angular pose.

In order to determine the rigid rotation between the finger sensor frames of references, $\{\mathcal{F}_k\}, k \in \{1,2,3,4,5\}$, and the palm frame of reference, $\{\mathcal{P}\}$, all sensors are used to measure the common vertical direction, as shown in fig. 2. When the sensed acceleration is equal in magnitude to gravity, the sensed direction is the vertical. The rigid rotation between the object frame of reference $\{\mathcal{O}\}$ and the palm is determined in the same way.

The distributed accelerometers provide a set of observed accelerations vectors $\mathbf{a}_s$, with $s \in \{0,1,2,3,4,5,obj\}$ (0-palm, 1-thumb, $2-5$-fingers, obj object). To deal with the non-observability of rotations about the vertical, a new observation of the same grasp at a distinct pose relative to the vertical is required, providing the set $\{\mathbf{a}_s|_t, \mathbf{a}_s|_{t+1}, \}$.

If n observations are made for distinct hand positions, recording the vertical reference provided by the inertial sensors, the absolute orientation can be determined using the orthogonal Procrustes method for 3D attitude estimation. We will use Horn's closed-form solution for absolute orientation using unit quaternions [15], applied here only to unit vectors. Since we are only observing a 3D direction in space, we can only determine the rotation between the two frames of reference. A single observation will not be enough to estimate the rotation, since rotations about the vertical direction are not observable, however a few observations with distinct hand orientations are enough. In [7] the same method is applied to the cross calibration of cameras and inertial sensors, and an error sensitivity analysis is performed with extensive simulation results, showing that even when not too diverse orientations are used, the geometric dilution of precision from such a narrow field of observation leads to the poorer results, but these improve with increasing number of observations. For the hand this means that if the hand is moving slightly when the grasp is stable, there is enough data to estimate the relative rotations.

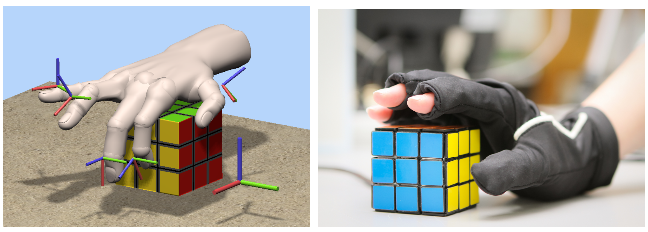


Fig. 2. Accelerometer sensor sensor axis distributed on the hand and object, and actual instrumented object and glove with the accelerometers.

Consider the case of finding the rotation between a single finger and the palm. Let ${}^{\mathcal{F}}\mathbf{v}_i$ be a measurement of the vertical by the inertial sensor on the finger, and ${}^{\mathcal{P}}\mathbf{v}_i$ the corresponding measurement made by the inertial sensor on the palm (from the sensor data, $\mathbf{v}_i = \mathbf{a}_s|_i / \|\mathbf{a}_s|_i\|$). We want to determine the unit quaternion $\mathring{\mathbf{q}}$ that rotates inertial measurements in the finger sensor frame of reference $\{\mathcal{F}\}$ to the palm sensor frame of reference $\{\mathcal{P}\}$. For the n observations, we want to find the unit quaternion $\mathring{\mathbf{q}}$ that maximises

$$\sum_{i=1}^{n} (\mathring{\mathbf{q}}\ {}^{\mathcal{F}}\mathbf{v}_i\ \mathring{\mathbf{q}}^*) \cdot {}^{\mathcal{P}}\mathbf{v}_i \tag{1}$$

After rearranging terms and some manipulation, using ${}^{\mathcal{F}}\mathbf{v}_i = ({}^{\mathcal{F}}x_i, {}^{\mathcal{F}}y_i, {}^{\mathcal{F}}z_i)^T$ and ${}^{\mathcal{P}}\mathbf{v}_i = ({}^{\mathcal{P}}x_i, {}^{\mathcal{P}}y_i, {}^{\mathcal{P}}z_i)^T$ this can be re-written as finding $\mathring{\mathbf{q}}$ that maximises $\mathring{\mathbf{q}}^T\, \mathbf{N}\, \mathring{\mathbf{q}}$ where

$$\mathbf{N} = \begin{bmatrix} (S_{xx}+S_{yy}+S_{zz}) & S_{yz}-S_{zy} & S_{zx}-S_{xz} & S_{xy}-S_{yx} \\ S_{yz}-S_{zy} & (S_{xx}-S_{yy}-S_{zz}) & S_{xy}+S_{yx} & S_{zx}+S_{xz} \\ S_{zx}-S_{xz} & S_{xy}+S_{yx} & (-S_{xx}+S_{yy}-S_{zz}) & S_{yz}+S_{zy} \\ S_{xy}-S_{yx} & S_{zx}+S_{xz} & S_{yz}+S_{zy} & (-S_{xx}-S_{yy}+S_{zz}) \end{bmatrix} \tag{2}$$

with

$$S_{xx} = \sum_{i=1}^{n} {}^{\mathcal{F}}x_i\, {}^{\mathcal{P}}x_i\ ,\ S_{xy} = \sum_{i=1}^{n} {}^{\mathcal{F}}x_i\, {}^{\mathcal{P}}y_i \tag{3}$$

and analogously for all 9 pairings of the components of the two vectors, matrix $\mathbf{N}$ can be expressed using these sums as in (2). The sums contain all the information that is required to find the solution. Since $\mathbf{N}$ is a symmetric matrix, the solution to this problem is the four-vector $\mathbf{q}_{max}$ corresponding to the largest eigenvalue λ_{max} of $\mathbf{N}$ - see [15] for details.

This is done for each of the fingers and the object, using the n observations, so that the angular pose relative to the palm can be determined. The above method finds the rotation that maximises the alignment of the rotated sensed verticals from the object and fingertip sensors to the palm sensed vertical, expressed by (1). The inertial frame verticals, ${}^{\mathcal{F}_k}\mathbf{v}_i$, ${}^{\mathcal{O}}\mathbf{v}_i$, and ${}^{\mathcal{P}}\mathbf{v}_i$, are easily obtained from the distributed accelerometers. The only restriction is that the system has to be motionless, or subject to constant speed, so that gravity can be used as a vertical reference.

With the closed form solution we obtain the relative rotation quaternions between each finger and the palm, and the object and the palm when using instrumented objects, $\mathring{\mathbf{q}}_1, \mathring{\mathbf{q}}_2, ..., \mathring{\mathbf{q}}_5, \mathring{\mathbf{q}}_{obj}$,. This is our feature space for identifying the grasp types. A given observation is compared against a library of grasp types in this feature space.

IV. Fitting Relative Pose of Fingers and Object Tactile Data to Known Grasp Types

For this work we are using the GRASP consortium comprehensive human grasp taxonomy presented in [12], a summary of which is shown in fig. 3. To confine the taxonomy to the goals of the GRASP project, grasp is defined as: *A grasp is every static hand pose with which an object can be held securely with one hand.* This

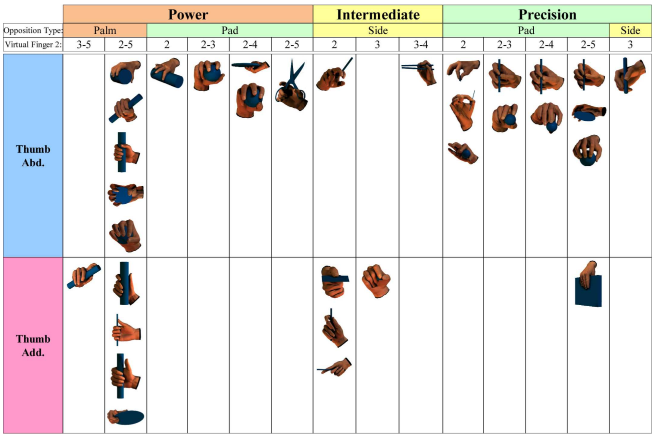


Fig. 3. Comprehensive Grasp Taxonomy which includes 33 different grasp types [12].

implies that the grasp stability has to be guaranteed irrespective of the relative force direction between object and hand. Therefore intrinsic movements are excluded because the object is not in a constant relationship to the hand. The use of both hands and gravity dependent grasps is also excluded. For instance, the Hook Grasp and the Flat Hand Grasp are not considered, since the hand extrinsic orientation is vital to the grasp stability.

In fig. 3 the top classification in columns is done by the power/precision requirements. The finer differentiation is done, depending on whether the opposition type is Palm, Pad or Side Opposition. The opposition type is also defining the virtual finger VF 1: In the case of Palm Opposition the Palm is mapped into VF 1, in Pad and Side Opposition the Thumb is VF 1. The only exception to this rule is the Adduction Grasp, where the thumb might not even contact with the object. To differentiate between the two rows, the position of the thumb is used. The classification here depends on whether the CMC joint of the thumb is in an adducted or abducted position [12]. The virtual finger (VF) grouping is used to define individual grasps, VF1 is defined by the opposition type and VF2 and VF3 by the fingers that together form the virtual finger.

Since many grasps have the same properties (opposition type, thumb position etc.), some cells in fig. 3 are populated with more than one grasp, the only difference being many times in the shape of the object. This offers the possibility to reduce the set of all 33 grasps down to 17 grasps by a merge of the grasps within one cell to a corresponding standard grasp. For our work we are only using the higher level 17 grasp types of the grasp classification [12].

As indicated in the previous section, from the distributed accelerometers we obtain the relative rotation quaternions between each finger and the palm, and the object and the palm when using instrumented objects, $\mathring{\mathbf{q}}_1, \mathring{\mathbf{q}}_2, ..., \mathring{\mathbf{q}}_5, \mathring{\mathbf{q}}_{obj}$,. This is our feature space for identifying the grasp types. A given observation is compared against a library of grasp types in this feature space. The distance metric used is the angular difference between the corresponding quaternions, given by

$$\delta\mathring{\mathbf{q}} = \mathring{\mathbf{q}}^{-1} * \mathring{\mathbf{q}}_{lib} \tag{4}$$

$$\theta_{\delta q} = 2\cos^{-1}(\delta q_s) \tag{5}$$

where $\delta q_s$is the scalar component of $\delta\mathring{\mathbf{q}}$. We take the absolute value $\delta_\theta = |\theta_{\delta q}|$ as the distance measure.

For the object we have to consider the symmetries, for instance for the cube, grasp-wise, any right angle rotation keeps the same orientation to the hand. So we have to compare with $\mathring{\mathbf{q}}_{obs}$ in 6 variations corresponding to distinct faces of the cube being adjacent to the hand, since the grasp type will be the same. For other objects this will also need to be taken into account.

The above feature space is not sufficient to fully determine the grasp type, since the angular pose, relative to the palm, of the fingers and object is similar in some of the grasps. By using instrumented objects, not only with accelerometers, but also with tactile sensors, we can use the number of contact sensing points or contact area and to some degree the intensity to narrow down the range of grasps types to be matched in the above feature space.

For the instrumented cube used (fig.2 and fig. 4) each face has 9 tactile sensors, so we have as tactile data $\mathbf{T}_c$ as a matrix of $t_{f,i}$ with $f \in \{r, g, y, b, o, w\}$ (coloured faces) and cell index within each face $i \in \{1, ..., 9\}$.

The number os cells above a minimum force threshold is used to distinguish between a power grasp and a precision grasp. We follow a decision tree: first the contact area determines if it is a power grasp or a precision grasp, then the nearest neighbour in the feature space of the relative finger-palm angular pose is found against the library of pre-recorded grasp types.

V. Experimental Results

Figure 4 shows the experimental setup used, and an example of a grasp being performed.

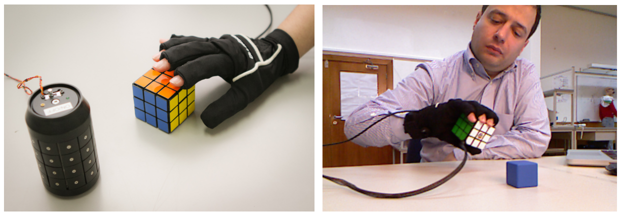


Fig. 4. Experimental setup used: a glove with distributed accelerometers (Acceleglove[5]), and instrumented objects, cube and soda can, with tactile and inertial sensors [3]; and example of performing a grasp.

The instrumented cube has one tri-axial accelerometer per face, but we consider a single one that averages de values from the 6 sensors taking into account the geometry. The soda can has 10 faces in a ring, again we consider a single tri-axial acceleration measure. The cube has 9 tactile cells per face, providing the measurement matrix $\mathbf{T}_c$ as indicated above, fig. 5 shows overlaid values.

To have a set verticals from acceleration measurements $\mathbf{a}_s$ described above, we consider that if the modulus of the sensed acceleration is close to gravity it can be used as a vertical reference (fig. 6). The accelerometers are sampled at 20 Hz, and when the modulus is below a certain threshold, a new observation for the relative pose is performed every 0.5 s (every 10 samples), with a local window of 3 samples to filter sensor and vibration noise. The cube has a higher sample rate (500 Hz) but we subsample to the glove sample rate.

To evaluate the quality of the estimated relative angular pose, we re-project the finger observations to the palm frame of reference,

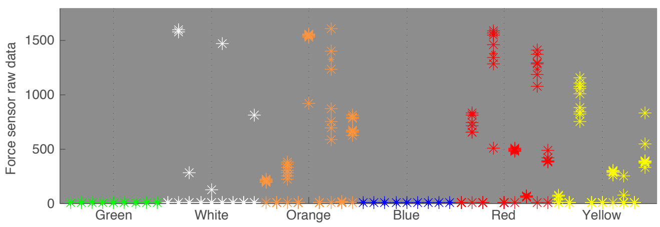


Fig. 5. Overlay of tactile force distribution on the instrumented cube during manipulation, 9 sensing cells per coloured face.

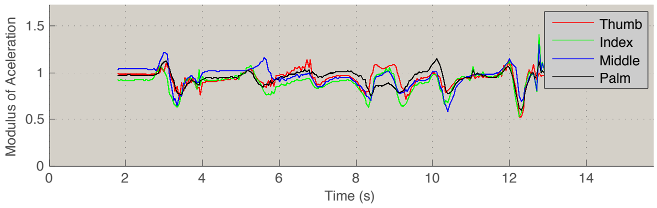


Fig. 6. Modulus of sensed acceleration vectors for sample point selection.

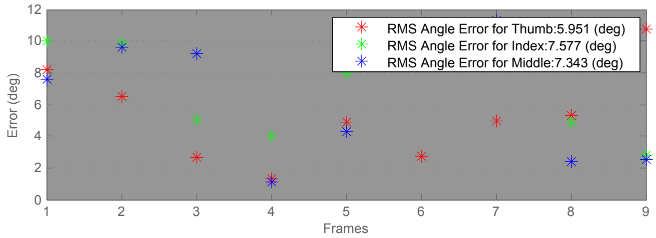


Fig. 7. Re-projection error of the estimated angular finger and object pose.

and take the angular mismatch as an error measure. From the results presented in fig. 7 we see that the r.m.s error can be up to 10 deg, but this is enough for our grasp type classification.

Tables I shows the grasp type identification, indicating the distances of the observed relative pose to the grasp types in the library, as well as the contribution of the tactile data to distinguish power grasps from precision grasps. The gesture being performed is a power/palm/2-5 grasp as shown in figure 4. It is also the nearest neighbour considering a Manhattan distance in the feature space, in this case about 27 deg.

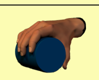
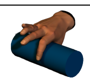
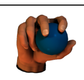
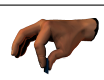
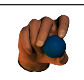
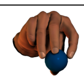
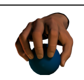
VI. Conclusions and Future Work

We presented a clutter free and easily deployable system of identifying human natural manipulation of objects. By using small accelerometers distributed on the back of the fingers and hand, and instrumenting the manipulated object, we are able to determine the grasp type and contact points with the object. Preliminary results show that the method works, although many aspects need to be worked upon to improve robustness and range of detected grasp types. Although for our experimental setup we used a commercial glove with accelerometers, we intend to build a custom one that can attach the minute sensors to the fingers and palm without a full glove. The used instrumented object, the cube, is wired, but we have under final development wireless versions including the soda can object shown in fig. 4. This will enable studying human manipulation in diverse environments and situations.

We intend to explore the use of more accelerometers, one on each finger segment, to enable the full reconstruction of the hand pose, making it more robust in identifying the grasp types. Going beyond identifying sequences of stable grasps, we also intend to further explore in-hand manipulation using the dynamic inertial data to classify intrinsic hand movements as proposed in [16]. Here intrinsic movements are defined as coordinated movements of the digits to manipulate and object within the hand. They are contrasted with extrinsic movements, defined as movements of a prehended object by displacement of the hand as a whole. The intrinsic movements are subdivided into simultaneous, exploring simple and reciprocal synergies, and sequential patterns. In order to deal with non-observability issues more accelerometers will be required and eventually gyrometers and magnectic sensors for the wrist or palm. The geometric distribution of contact points provided by the tactile sensing on the object can also help to overcome non-observability issues.

TABLE I

Result of grasp type identification

Op : VF	power Pal : 2-5	power Pad: 2	power Pad: 2-3	power Pad: 2-4	prec. Pad: 2	prec. Pad: 2-3	prec. Pad: 2-5	prec. Pad: 2-5	power Pal: 2-5	int Side: 2	prec. Pad: 2-5
tactile cell cnt:	18	13	11	10	6	6	8	9	11	2	2
											
m.. dist. (deg):	**27**	51	53	61	31	32	33	44	136	207	56

Distance of the observed grasp to the grasps in the library, and masking of range of grasps using the tactile data. The identified grasp type is the power/palm/2-5 grasp, corresponding to the lowest value of the distance, the cell count of the trial was 17, indicating a power grasp.

References

[1] L. Dipietro, A.M. Sabatini, and P. Dario. A survey of glove-based systems and their applications. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(4):461 –482, july 2008.

[2] J.K. Perng, B. Fisher, S. Hollar, and K.S.J. Pister. Acceleration sensing glove (asg). In *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*, pages 178 –180, 1999.

[3] HANDLE project, ICT-236410, www.handleproject.eu, 2011.

[4] N. Hendrich, D. Klimentjew, and Jianwei Zhang. Multi-sensor based segmentation of human manipulation tasks. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2010 IEEE Conference on*, pages 223 –229, 2010.

[5] J.L. Hernandez-Rebollar, N. Kyriakopoulos, and R.W. Lindeman. The AcceleGlove: A Whole-Hand Input Device for Virtual Reality. In *(Technical Sketch), Conference Abstracts and Applications, ACM SIGGRAPH*, page 259, July 2002.

[6] Pedro Trindade and Jorge Lobo. Distributed accelerometers for gesture recognition and visualization. In *DoCEIS'11 - Doctoral Conference on Computing, Electrical and Industrial Systems*, pages 215–223, Lisbon, Portugal, February 2011.

[7] Jorge Lobo and Jorge Dias. Relative pose calibration between visual and inertial sensors. *The International Journal of Robotics Research (IJRR) Special Issue from the 2nd Workshop on Integration of Vision and Inertial Sensors.*, 26:561–577, 2007.

[8] Jorge Lobo and Jorge Dias. Vision and inertial sensor cooperation using gravity as a vertical reference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1597–1608, December 2003.

[9] M.R. Cutkosky. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *Robotics and Automation, IEEE Transactions on*, 5(3):269 –279, June 1989.

[10] J. R. Napier. The prehensile movements of the human hand. *J Bone Joint Surg Br*, 38-B(4):902–913, 1956.

[11] Onno A. van Nierop, Aadjan van der Helm, Kees J. Overbeeke, and Tom J. P. Djajadiningrat. A natural human hand model. *Vis. Comput.*, 24:31–44, November 2007.

[12] T. Feix, R. Pawlik, H.-B. Schmiedmayer, J. Romero, and D. Kragic. The generation of a comprehensive grasp taxonomy. In *Robotics, Science and Systems Conference: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation*, June 2009.

[13] GRASP project, Humangraspingdatabase, http://web.student.tuwien.ac.at/ e0227312/, 2011.

[14] Javier Romero, Thomas Feix, Hedwig Kjellstrm, and Danica Kragic. Spatio-temporal modeling of grasping actions. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipeh,Taiwan, October 2010.

[15] B.K.P Horn. Closed-Form Solution of Absolute Orientation Using Unit Quaternions. *Journal of the Optical Society of America*, 4(4):629–462, April 1987.

[16] J. M. Elliott and K. J. Connolly. A classification of manipulative hand movements. *Developmental Medicine & Child Neurology*, 26(3):283–296, 1984.