

# Making Blind People Autonomous in the Exploration of Tactile Models: A Feasibility Study

Francesco Buonamici, Rocco Furferi<sup>(✉)</sup>, Lapo Governi,  
and Yary Volpe

Department of Industrial Engineering of Florence, University of Florence (Italy),  
Via Di Santa Marta 3, 50139 Florence, Italy  
{francesco.buonamici, rocco.furferi, lapo.governi, yary.  
volpe}@unifi.it

**Abstract.** Blind people are typically excluded from equal access to the world's visual culture, thus being often unable to achieve concrete benefits of art education and enjoyment. This is particularly true when dealing with paintings due to their bi-dimensional nature impossible to be explored using the sense of touch. This may be partially overcome by translating paintings into tactile bas-reliefs. However, evidence from recent studies suggests that the mere tactile exploration is often not sufficient to fully understand and enjoy bas-reliefs. The integration of different sensorial stimuli proves to dramatically enrich the haptic exploration. Moreover, granting blind people the possibility of autonomously accessing and enjoying pictorial works of art, is undoubtedly a good strategy to enrich their exploration. Accordingly, the main aim of the present work is to assess the feasibility of a new system consisting of a physical bas-relief, a vision system tracking the blind user's hands during "exploration" and an audio system providing verbal descriptions. The study, supported by preliminary tests, demonstrates the effectiveness of such an approach capable to transform a frustrating, bewildering and negative experience (i.e. the mere tactile exploration) into one that is liberating, fulfilling, stimulating and fun.

**Keywords:** Cultural heritage · Blind people · Hand tracking · Human-computer interaction

## 1 Introduction

Promoting, protecting and ensuring the full and equal enjoyment of all human rights and fundamental freedoms by all persons with disabilities is a worldwide high priority issue. Enjoying the Cultural Heritage makes no exception! This is particularly true for blind people (BP) since they are inevitably disadvantaged in enjoying artworks usually created for sighted people. In fact, for BP, the accessibility to cultural or artistic areas is affected not only by mobility impairment but rather by the inability to sense the artworks themselves. Moreover, when dealing with paintings, the inherently bi-dimensional structure of such artworks considerably complicates the experience of BP, practically excluding them in enjoying such kind of art.

This may be partially overcome by translating paintings into 3D models i.e. into tactile bas-reliefs to be explored by using the sense of touch. Not coincidentally, renowned institutions such as the Omero Tactile Museum (Bologna), the Art Institute of Chicago and the Tactile Gallery at the Wellcome Gallery of London, created a collection of famous paintings translated in the form bas-relief-like representations, oriented to the aesthetic education of blind people.

Unfortunately, the manufacture of tactile bas-reliefs is a complex task and, to date, entails the work of appositely trained artists (sculptors), thereby drastically increasing the cost and reducing both the production rate and the total number of available reproductions. With the aim of speeding up the manufacture of tactile bas-reliefs, in recent years a number of significant studies have been carried out making use of Computer-Aided techniques dealing with bas-relief reconstruction from images [1–3].

Authors themselves, recently, introduced a number of Computer Aided techniques for semi-automatic translation of paintings appositely devised to improve blind people's accessibility to bi-dimensional artworks [4–8]. The developed methodology, integrating perspective-based scene reconstruction with shape from shading-based shape retrieval, allows the virtual bas-relief reconstruction that can be converted into a physical tactile model possibly by means of Rapid Prototyping techniques (see Fig. 1).

Even though the scientific literature is moving towards the development of high quality 3D models from paintings, evidence from recent studies performed with blind people accessing this kind of reproductions suggests that the mere tactile exploration is often not sufficient to fully understand and enjoy them; therefore blind and visually impaired people often require to be specifically (though briefly) trained and assisted by a sighted accompanying person [9, 10] (see Fig. 2a). This is often perceived as a limitation for the blind who increasingly demand for autonomous accessibility to the world of cultural heritage.

From this point of view, the use of a hand-tracking system able to determine which area of the bas-relief is touched and to provide a real-time verbal description and/or other kinds of audio feedback could dramatically improve the blind people understanding and enjoyment during un-assisted exploration of artworks.



**Fig. 1.** From painting (“Madonna with Child and Angels” by Niccolò Gerini di Pietro, Villa La Quiete, Firenze, Italy) to bas-relief using the semiautomatic procedure developed by the authors.



**Fig. 2.** (a) Typical exploration of artworks where skilled persons guide the hands of blind people [11] – (b) Conceptual framework of the exploration performed by the supposed system.

Accordingly, the main aim of the present work is to assess the feasibility of a new system consisting of a physical bas-relief (obtained using for instance authors’ 3D modeling approach), a vision system tracking the blind user’s hands during “exploration” and an audio system providing verbal descriptions (see Fig. 2b).

Such a novel system requires at least (1) a 3D acquisition device + software package to track the user hands; (2) a number of algorithms capable of detecting the position of the bas-relief in the same reference frame defined by the acquisition sensor; (3) a number of algorithms aiming at detecting the position and the distance of the user hand/finger with respect to the model; (4) the complete knowledge of the digital 3D bas-relief model and (5) an appropriate verbal description linked to relevant objects/subjects in the scene. As a consequence, the definition of a system capable of assisting BP in autonomous exploration of a bas-relief should take into account relevant issues such as hand tracking, point cloud registration and distance evaluation between user’s fingertip and bas-relief model.

## 2 Background

Since, as mentioned in the introductory section (and deeply investigated in the Sect. 3), the proposed system encompasses a number of methods for hand tracking, point cloud registration and distance evaluation, a brief reminder of the current state of the art related to these issues is provided below.

### 2.1 Hand Tracking

The term “hand tracking” encloses a number of techniques of the “human motion capture” family sharing with them the common purpose of locating the position of a human hand. This area of study offers massive possibilities to applications in the field of human-computer interaction [16], for example gestural interfaces to improve easiness of use of computer systems, or to interact with virtual environments, as in [17, 19, 20].

Furthermore, hand tracking techniques are widely implemented to help impaired people with everyday life in multiple situations [18, 21, 22].

The problem of tracking a human hand is tackled in literature through different approaches, which distinguish themselves by the hardware or the algorithm used. Wearable devices, visual markers, electromechanical sensors and actuators are some examples of the hardware used in various strategies.

The focus of the present work is on vision-based techniques, using an optic sensor observing the scene in order to get the data required to perform the tracking. In fact they allow maximum haptic sensitivity and gestural freedom to the user (essential requisite for BP).

Vision-based techniques can be classified into appearance - and model-based approaches [16, 23].

Appearance-based methods typically establish a mapping from a set of image features to a discrete, finite set of hand model configurations [12]. Such methods, as described in [18, 24], are computationally efficient and “well suited for problems such as hand posture recognition where a small set of known target hand configurations needs to be recognized. Conversely, such methods are less suited for problems that require an accurate estimation of the pose of freely performing hands” [12].

Model-based approaches [12, 24] use a digital model of a hand to simulate and estimate the position of the user hand in the observed scene. This is usually performed by solving an optimization problem whose objective function measures the discrepancy between the visual cues that are expected due to a model hypothesis and the actual ones. Even if these methods require costly algorithms (and accordingly the achievement of satisfying frame rate values during tracking is reached only by means of high-end computer hardware) they prove to be effective for a more refined estimation of hand position. Consequently, in this work model-based methods are investigated with particular reference to the model proposed by [12] and further developed in [25–27]. In such a model, 3D data provided by a Microsoft Kinect® are used to isolate the user hand in 2D and 3D by means of a skin color detection followed by depth segmentation. The hand model (palm and five fingers) is described by geometric primitives and parametrized encoding 26-DOF (i.e. is represented by 27 parameters). The optimization procedure is carried out by means of a Particle Swarm Optimization technique [13]. The procedure contemplates temporal continuity of subsequent frames, searching for a solution in the neighborhood of the one found for the last frame analyzed. Some examples of the results obtained with this method are shown in Fig. 3.

## 2.2 Point Cloud Registration

Point cloud registration techniques aim at minimizing the distance between two given sets of points usually by performing a rigid roto-translation.

According to [28], where a review of the most relevant registration methods is provided, two distinct consequent phases are usually carried out: coarse registration and fine registration.

Coarse registration performs a rough estimation of the alignment between the two point clouds. Typically, this is achieved through iteration algorithms aiming at



**Fig. 3.** Snapshots from a hand tracking sequence from [12], the hand model in the hypothesized solution is placed upon the real hand RGB image taken by Kinect® (Color figure online).

matching common features in the two point clouds in order to obtain a set of correspondences. A number of algorithms can be used to perform the coarse registration: Point Signature, Spin Image, RANSAC-based (RANdom SAMple Consensus), PCA (Principal Component Analysis) and genetic algorithms [28]. As described in the next section, in the present work coarse registration is performed by means of an appositely devised interactive procedure based on the hand tracking system.

Fine registration performs a more accurate solution by using the coarse registration as an initial guess. In recent years, some methods have been presented in scientific literature: (a) iterative closest point (ICP); (b) Chen’s method; (c) signed distance fields; and (d) genetic algorithms, among others [28].

Devised by Besl & McKay [29], the ICP method minimizes the distance between point-correspondences (i.e. closest points). Let  $A$  and  $B$  be the two point sets to be aligned; for each of the  $N_A$  points  $a_i$  belonging to the point cloud  $A$  the closest point  $b_i$  in the point cloud  $B$  is found using nearest neighbor search techniques (see Sect. 2.3). Using a mean squared error cost function (see Eq. 1) the combination of rotation (matrix  $R$ ) and translation (vector  $t$ ) that best align each point  $a_i$  to its match found in  $B$  is then computed.

$$f = \frac{1}{N_A} \sum_{i=1}^{N_A} \left\| \vec{a}_i - (R \cdot \vec{b}_i + t) \right\|^2 \quad (1)$$

This process is iterated, using the new set of point  $A$  rotated and translated according to  $R$  and  $t$ , until a convergence criterion is satisfied.

Differently from ICP approach, Chen’s method uses point-to-plane distance in the optimization (rather than point-to-point distance). Distance-to-plane evaluation is more laborious, but the algorithm is more robust to local minima and less sensible in the presence of non-overlapping regions [28] thus allowing to obtain better results. As described in Sect. 3, in the present work fine registration is performed by implementing both methods.

### 2.3 Distance Evaluation

One of the more known methods to evaluate a distance between two 3D points is the Nearest Neighbor algorithm, an optimization problem for finding closest (or most similar) points. Among the huge number of methods proposed in literature, the most

frequently used and reliable one is the so called “KD Tree” [14]. This approach relies upon an efficient organization of the data set; a tree-like structure is built to organize points in a k-dimensional space thus allowing to quickly computing the “nearest neighbor” of a given point.

In particular, KD Tree performs recursive binary partitions of the dataset (tree structure), creating regions circumscribed by K-dimensional hyper-planes. In a 3D tree the planes are usually created following the axis directions, allowing for particularly convenient computation of the distances. The points are, therefore, enclosed in well-defined regions that are usefully mapped to avoid redundancy in the computation of the nearest neighbor.

### 3 Feasibility Analysis

In this section the development of the prototypal system, sketched in Fig. 2a, is discussed. The main effort has been to build a first working prototype (although rudimentary) of the system to help BP in enjoying artworks. The effectiveness of the prototype would demonstrate the feasibility of such a system to be, once improved, installed in museums in the next future.

For developing the feasibility study, as already mentioned, the starting point consists of a high resolution digital 3D bas-relief and its physical counterpart. By a way of example, in the following description the tactile reproduction of the “Madonna with Child and Angels” of Niccolò Gerini di Pietro (in exposition at “Villa la Quiete”, Firenze, Italy) is used to explain the overall process (see Fig. 1). The digital model, created using the approach described in [8], has a resolution of about 0.17 mm per pixel (depending on the resolution of the acquired image); the physical prototype is sized approximately 87 mm × 47 mm × 9 mm.

Let, accordingly, assume that the physical prototype (1) is conveniently arranged to a support allowing the user to easily touch it (e.g. bas-relief positioned at 0.8 m height and tilted by 20° with respect to the horizontal plane) and (2) the bas-relief falls inside a 3D acquisition device field of view.

With reference to the 3D acquisition device, in this feasibility analysis a Microsoft Kinect® is chosen as preferred hardware. As widely known, such a device consists of a projector-camera triangulation sensor able to obtain a streaming of 3D data of the observed scene with a maximum resolution of 1.3 mm per pixel inside an angular field of view of 57° horizontally and 43° vertically and through a range of approximately 0.7-6 m. The preferred frame rate used for the acquisition is in the range 15-20 fps. Frame rate is a key factor in hand tracking, and is indeed directly related to the quality of tracking achievable from the point of view of both stability and hand movement speed traceable by the method.

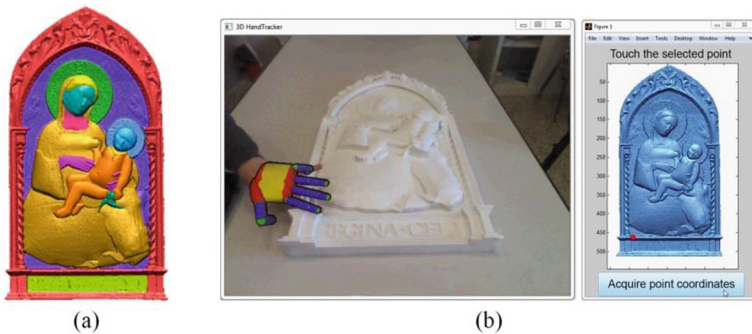
To track user’s hand movement, the method proposed in [12] is adopted and the “Hand Tracking Library” proposed in [15] is used as a basis for implementing the proposed approach. This choice is motivated by the good performances obtained during introductory tests carried out by authors of the present paper. Such tests demonstrated both good overall performances (in terms of frame rate, accuracy and stability) and consistent results with reference to data available in literature.

Hand tracking is limited in this work, due to its conceptually nature, only to the right index fingertip; therefore, a custom algorithm to trace its position in the Kinect® reference system, based on the hand tracking library data, was implemented into MATLAB® environment. Moreover, to optimize hand tracking the 3D sensor has been placed at a distance of approximately 1 m from the hand, according to the suggestions provided in [12] and preliminary tests carried out by the authors.

Once the position of the user’s finger is tracked in real time, it’s necessary to know the position of the bas-relief relatively to the same reference system. This problem is tackled taking advantage of the Kinect®, using it as a 3D scanner to obtain, with a single placement, a 3D model of the bas-relief correctly referenced to the Kinect® reference frame system. This procedure allows the detection of the forefinger position with respect to the digital and physical models at the same time. This is a prime issue in order to identify what the user is touching and to provide the necessary information.

Unfortunately, Kinect® sensor provides low definition (LD) and highly noisy scans which are not optimal to be used in the subsequent phase where the contact between the user forefinger and the physical model needs to be precisely detected. For this reason the (available) high definition (HD) virtual model of the bas-relief is used in order to retrieve more accurate information. In other words, if on one side the LD model is useful to refer the physical prototype to the user finger (in the Kinect reference frame), the availability of the HD model is crucial for a better discrimination of the touched area. Moreover, since each significant area (e.g. the face of Virgin Mary in Fig. 4.(a) of the model has to be associated to a verbal description to be provided to the user, the HD model should be appropriately segmented. The 3D segmentation of the model, whose description falls outside the scope of the present paper, can be practically performed using dedicated software packages such as Rapidform® and the choice of the number of segments strictly depends on the information to be delivered to the user.

To refer the HD point cloud to the Kinect reference frame, a point cloud registration of the HD and LD models is required. Such a registration process, as mentioned in the previous section, is achieved in two steps: first, a coarse registration is carried out to roughly align HD and LD point clouds; then, a fine registration allows to properly refer the HD model to the reference frame.



**Fig. 4.** (a) Segmented 3D digital model of the “Madonna with Child and Angels”, by Niccolò Gerini di Pietro. – (b) GUI for coarse registration between the HD and LD models.

Coarse registration is performed using an interactive “custom-built” approach that takes advantage of the hand tracking system. The user is asked to touch with the tracked finger a number  $N$  (e.g. 6) of non-aligned areas on the physical bas-relief; the  $N \times 3$  matrix of coordinates  $C_{LD}$  locating the  $N$  finger-bas-relief contact points are then memorized.

Subsequently, using an appositely devised GUI, the user is required to select the HD model points roughly corresponding to the ones actually touched in the physical model. The result of this phase is to retrieve a matrix of coordinates  $C_{HD}$  describing the  $N$  points in the HD model (see Fig. 4b).

The general relationship between matrices  $C_{LD}$  and  $C_{HD}$  is described as follows:

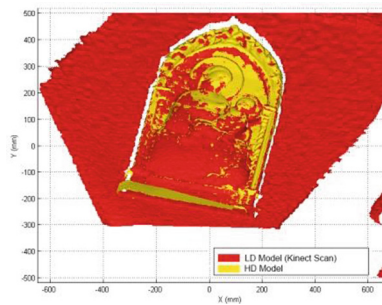
$$C_{LD} = R * C_{HD} + t \quad (2)$$

To achieve the coarse registration, the rotational matrix  $R$  and translational vector  $t$  have to be evaluated from Eq. 2. This may be easily done using a SVD-based procedure [29, 31].

The coarse registration is used as an initial guess for the subsequent fine registration. In order to identify the best method (among the ones identified at the state of the art) to perform such a fine registration, a number of tests have been carried out implementing both the ICP and Chen’s algorithms.

Despite according to literature Chen’s method is usually most stable with respect to ICP one, the use of such algorithm to this specific application lead to unsatisfactory registration results. On the other hand, the ICP method turned out as the most reliable, although it proves to be less computationally efficient. Average time for convergence of ICP is in the range of 5–8 min with a RMS error between 2–3 mm; such an error value is comparable with the Kinect scan accuracy. On the basis of the above mentioned tests, the ICP method is used for fine registration. An example of final results obtained after the registration phase is depicted in Fig. 5.

Once the registration phase is performed, a real time computation of the distances between the user finger and the points of the HD model (correctly referenced in the Kinect frame system) is required. This step is carried out using the KD Tree algorithm into MATLAB® environment with increasing dataset dimension and increasing query



**Fig. 5.** LD and HD models of the “Madonna with Child and Angels” by Niccolò Gerini di Pietro after coarse and fine registration.



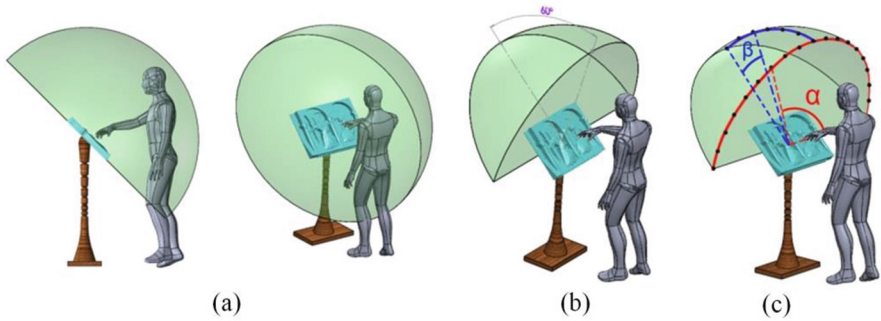
point number (this is made to forecast future implementations using information from multiple fingers touching the model at the same time). The results of the application of such an algorithm are compared, in terms of computational time, with other known methods such as N-D nearest point search [30] and brute force.

KD Tree proves to be the faster method, performing at about 0.02 s for a dataset comparable with the model dimensions and with a single query point. This means that the distance evaluation is performed with a 50 Hz frequency that is higher than the hand tracking fps acquisition (equal to maximum 20 fps). In evaluating the computational time it has to be noticed that the time required to create the tree-structure for performing the KD-tree algorithm is in an “offline” operation and, therefore, does not affect the real-time evaluation.

Finally, for every frame, the distance of the nearest neighbor to the query point (the right index fingertip) is compared to a threshold value (empirically set) that establishes the fulfillment of the touching condition. If the distance is smaller than the threshold, an algorithm searches for the segmented area in the 3D HD model corresponding to the actually touched region of the physical bas-relief. Since each segment is associated to a file containing the desired verbal description, this procedure allows the user to retrieve the information about the touched area.

The description provided above proves the feasibility from a technical point of view taking into account how to merge together all frameworks (hand tracking, registration, distance evaluation, etc.) to build an assistive system for BP. Obviously, in the hypothesis of a potential museum collocation of an improved system, the study of a possible final layout is required in order to maximize technical performances of the system together with accessibility and easiness for BP.

One of the most relevant parameters affecting the hand tracking performance is the sensor position; therefore, some tests are carried out to analyze the influence of the Kinect® position on the performance of the hand tracking algorithm, taking into account also the possible optical occlusions during the user-exploration phase. Prior to define the optimal position of the 3D sensor, it is necessary to decide the final disposition of the bas-relief to be explored by BP. To maximize the user ergonomics during tactile exploration (hypothetically placed into a museum environment), a preferred solution consists of positioning the bas-relief on a plane with a 45° inclination (with respect to the horizontal plane) and a height of approximately 1.2 m. As mentioned above, in order to maximize the hand tracking system performance, the sensor can be located in any point lying into a hemisphere with radius equal to 1 m (see Fig. 6a) centered in the bas-relief barycenter. Actually, from a practical point of view, this area can be reduced to a spherical sector with bevel equal to 60° (see Fig. 6b). In fact, this configuration allows reducing possible occlusions made by the user. On the basis of a series of tests performed using 4 different bas-reliefs, good hand tracking performance is obtained when the sensor is positioned with an angle  $\beta$  equal to 40° (see Fig. 6c). This is probably due to the fact that such a position allows the Kinect® to “observe” the hand mostly perpendicularly to the hand plane. Quite the reverse, positions with low  $\beta$  values lead to low stability. The configuration shown in Fig. 6c allows also to perform a virtual analysis of visible bas-relief areas. For each possible position described by varying angles  $\alpha$  and  $\beta$ , the visible percentage of the bas-relief is



**Fig. 6.** Possible collocation surface of the sensor: (a) hemisphere with radius 1 m. (b) spherical sector with bevel of  $60^\circ$  (considering user-caused occlusions). – (c)  $\alpha$  &  $\beta$  reference system for the collocation surface.

evaluated. The position with  $\alpha = 0^\circ$  and  $\beta = 50^\circ$  results the best solution since it provides averagely for the 4 case studies analyzed a visibility equal to 86 %.

Considering the performance of the hand tracking system together with the visibility analysis, the ideal position for the sensor turns out to be the one with  $\alpha = 0^\circ$  and  $\beta = 40^\circ$  corresponding to a visibility equal to 73 %.

## 4 Conclusion

In this work, the feasibility study of a new system consisting of a physical bas-relief, a vision system tracking the blind user's hands during "exploration" and an audio system providing verbal descriptions was provided. The hand tracking issue was tackled using the approach provided by [12], so that the position of the forefinger during the exploration of a tactile bas-relief is tracked in real time using Microsoft Kinect®. Then, the device was used as a 3D scanner to build a rough 3D model of the bas-relief correctly referenced to the Kinect reference frame system.

Since Kinect® sensor provides low definition and highly noisy scans which are not suitable to be used in the subsequent phase (where the contact between the forefinger and the model needs to be detected) the use of a high definition (HD) and less noisy virtual model was suggested. Being such a model already available as an outcome of author's previous works [8], an appositely devised procedure was used in order to register the HD model with the LD one.

Afterwards, using the KD Tree algorithm the region of the HD bas-relief nearest to the forefinger was determined so that the corresponding verbal description of the subject/object can be provided. Finally, the study of a possible final layout was performed in order to maximize technical performances of the system together with accessibility and easiness for BP.

Early tests demonstrated that the conceptual layout of the systems is quite sound even if a few limitations still subsist mainly due to the robustness of the tracking system which, from time to time, loses the target (hand) when it is moving too fast.

On the basis of preliminary tests performed with the support of a panel of BP, the proposed method results to be a first useful attempt to transform a frustrating, bewildering and negative experience (i.e. the mere tactile exploration) into one that is liberating, fulfilling, stimulating and fun.

Future work will be addressed to increase the performance and the robustness of hand tracking system (using for instance Kinect® 2.0 together) and to take into account multiple fingers touching the model at the same time.

## References

1. Li, Z., Wang, S., Yu, J., Ma, K.L.: Restoration of brick and stone relief from single rubbing images. *IEEE Trans. Vis. Comput. Graph.* **18**(2), 177–187 (2012)
2. Reichinger, A., Maierhofer, S., Purgathofer, W.: High-quality tactile paintings. *J. Comput. Cult. Heritage (JOCCH)* **4**(2), 1–13 (2011)
3. Reichinger, A., Neumüller, M., Rist, F., Maierhofer, S., Purgathofer, W.: Computer-aided design of tactile models. In: Miesenberger, K., Karshmer, A., Penaz, P., Zagler, W. (eds.) *ICCHP 2012, Part II. LNCS*, vol. 7383, pp. 497–504. Springer, Heidelberg (2012)
4. Governi, L., Furferi, R., Volpe, Y., Puggelli, L., Vanni, N.: Tactile exploration of paintings: an interactive procedure for the reconstruction of 2.5D models. In: *Proceeding of 2014 22nd Mediterranean Conference on Control and Automation (MED)*, Palermo, Italy (2014)
5. Volpe, Y., Furferi, R., Governi, L., Tennirelli, G.: Computer-based methodologies for semi-automatic 3D model generation from paintings. *Int. J. Comput. Aided Eng. Technol.* **6**(1), 88–112 (2014)
6. Governi, L., Furferi, R., Volpe, Y., Vanni, N.: Tactile 3D bas-relief from single-point perspective paintings: a computer based method. *J. Inf. Comput. Sci.* **11**(16), 1–14 (2014)
7. Governi, L., Carfagni, M., Furferi, R., Puggelli, L., Volpe, Y.: Digital bas-relief design: a novel shape from shading-based method. *Comput.-Aided Des. Appl.* **11**(2), 153–164 (2014)
8. Governi, L., Furferi, R., Volpe, Y., Carfagni, M., Puggelli, L., Vanni, N.: From 2D to 2.5D i.e. from painting to tactile model. *Graph. Models* **76**(6), 706–723 (2014)
9. Hayhoe, S.: An enquiry into passive and active exclusion from sensory aesthetics in museums and on the Web: two case studies of final year students at California school for the blind studying art works through galleries and on the web. *Brit. J. Visual Impairment* **32**(1), 44–58 (2014)
10. Hayhoe, S.: *Arts, Culture and Blindness: Studies of Blind Students in the Visual Arts*. Teneo Press, Youngstown (2008)
11. [http://www.cavazza.it/sites/default/files/images/immagini\\_museo/fotoletturatattile1.jpg](http://www.cavazza.it/sites/default/files/images/immagini_museo/fotoletturatattile1.jpg)
12. Oikonomidis, I., Kyriazis, N., Argyros, A.A.: Efficient Model-Based 3d Tracking of Hand Articulations Using Kinect. *BMVC*, Dundee (2011)
13. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: *International Conference on Neural Networks*, vol. 4, pp. 1942–1948. IEEE (1995)
14. Bentley, J.L.: Multidimensional binary search trees used for associative searching. *Comm. ACM* **18**(9), 509–517 (1975)
15. <http://cvrlcode.ics.forth.gr/handtracking/>
16. Pavlovic, V.I., Sharma, R., Huang, T.S.: Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 677–695 (1997)

17. Malik, S., Laszlo, J.: Visual touchpad: A two-handed gestural input device. In: ICMI 2004 - Sixth International Conference on Multimodal Interfaces, pp. 289–296 (2004)
18. Imagawa, K., Lu, S., Igi, S.: Color-based hands tracking system for sign language recognition. In: Proceedings - 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 462–467 (1998)
19. Wang, R.Y., Paris, S., Popovic, J.: 6D hands: Markerless hand tracking for computer aided design. In: UIST 2011 - Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, pp. 549–557 (2011)
20. Menelas, B., Hu, Y., Lahamy, H., Lichti, D.: Haptic and gesture-based interactions for manipulating geological datasets. In: Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, pp. 2051–2055 (2011)
21. González-Ortega, D., Díaz-Pernas, F.J., Martínez-Zarzuela, M., Antón-Rodríguez, M.: A Kinect-based system for cognitive rehabilitation exercises monitoring. *Comput. Methods Programs Biomed.* **113**(2), 620–631 (2014)
22. Boccanfuso, L., O’Kane, J.M.: CHARLIE: an adaptive robot design with hand and face tracking for use in autism therapy. *Int. J. Soc. Robot.* **3**(4), 337–347 (2011)
23. Erol, A., Bebis, G., Niclescu, M., Boyle, R.D., Twombly, X.: Vision-based hand pose estimation: a review. *Comput. Vis. Image Underst.* **108**(1–2), 52–73 (2007)
24. De La Gorce, M., Paragios, N., Fleet, D.J.: Model-based Hand Tracking With Texture, Shading and Self-occlusions. In: CVPR, pp. 1–8, June 2008
25. Oikonomidis, I., Kyriazis, N., Argyros, A.A.: Markerless and efficient 26-DOF hand pose recovery. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010, Part III. LNCS, vol. 6494, pp. 744–757. Springer, Heidelberg (2011)
26. Oikonomidis, I., Kyriazis, N., Argyros, A.A.: Full DOF tracking of a hand interacting with an object by modeling occlusions and physical constraints. In: ICCV, pp. 2088–2095, November 2011
27. Oikonomidis, I., Kyriazis, N., Argyros, A.A.: Tracking the articulated motion of two strongly interacting hands. In: Proceedings of IEEE Conference Computer Vision Pattern Recognition, pp. 1862 – 1869 (2012)
28. Salvi, J., Matabosch, C., Fofi, D., Forest, J.: A review of recent range image registration methods with accuracy evaluation. *IVC* **25**, 578–596 (2007)
29. Besl, P., McKay, N.: A method for registration of 3-d shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992)
30. Barber, C.B., Dobkin, D.P., Huhdanpaa, H.T.: The Quickhull Algorithm for Convex Hulls. *ACM Trans. Math. Softw.* **22**(4), 469–483 (1996)
31. [http://ngghiaho.com/?page\\_id=671](http://ngghiaho.com/?page_id=671)