ORIGINAL PAPER

Haptic Augmented Reality (HapticAR) for assembly guidance

J. C. Arbeláez[1] · **Roberto Viganò[1]** · **Gilberto Osorio-Gómez[2]**

Abstract

The use of Augmented Reality (AR) to support assembly tasks has been an area of interest from its origins in the 90s. Since then, the benefits that this technology could bring to assembly-related tasks have been shown. And, although several advances have been done in different areas such as software, hardware, and human interaction, there are still some problems that have not allowed AR to expand and reach its full potential. Thereby, authors propose a real-time vibrotactile guidance method based on the Gestalt continuity principle and developed a Haptic Augmented Reality application with a low-cost configuration to evaluate the support of the proposed method in assembly tasks. Thus, it potentially overcomes the existing visual issues of AR allowing the user to focus on the task and avoid over-reliance into the technology. The proposed system recognizes the different parts and sub-assemblies, generates the instructions to perform the assembly based on the target position and rotation of each part and verifies the assembly. Additionally, a test was conducted to guide the user in positioning a part, obtaining a high accuracy of rotation and location placement. Also, different functions of the application were tested and the results are suitable for supporting the user guidance.

Keywords Augmented Reality · Haptic · Assembly guidance · User interaction · Object recognition

1 Introduction

1.1 General issues

Augmented Reality (AR) dates from the 90s with the development of a prototype superimposing virtual information over the real world and using a see-through head-mounted display combined with a tracking and registration system. This technology was aimed to support some wiring activities at Boeing and had shown potential benefits in terms of efficiency and costs reduction [9].

Since then, several industrial applications have been reported over the past years showing the potential benefits of the use of such technology. Nevertheless, although there had been several advances in the enabling technology,[1] there are still many unsolved issues that overshadow its benefits, preventing its wide-spreading [13,28,36].

One of the major interest fields for the application of AR in the industry has been the assembly process. Where AR has been proposed as a support system for an assembly task recognizing the assembly parts, displaying related assembly instructions and verifying the operation [56].

Additionally, Augmented Reality and haptic technology have been used for enabling a user to feel a real environment enriched with a haptic stimulus, where it is possible to touch both virtual and real objects or real objects augmented with virtual touch [24].

The assembly guidance through AR has demonstrated some benefits compared with classic methods (digital and paper manuals) such as error rate reduction and a decrease of mental and physical effort. Displaying directly the information to the user, it is possible to avoid the attention swapping, the execution of repetitive movements and, at the same time, simplifying user's decisions [19,36,48,61].

✉ J. C. Arbeláez
   juancarlos.arbelaez@polimi.it

   Roberto Viganò
   roberto.vigano@polimi.it

   Gilberto Osorio-Gómez
   gosoriog@eafit.edu.co

[1] Department of Mechanical Engineering, Politecnico di Milano, Via G. La Masa, 1, 20156 Milan, Italy

[2] Design Engineering Research Group, Universidad EAFIT, Carrera 49 No 7 Sur-50, Medellín, Colombia

---

[1] Technologies required to develop an AR environment, related to software and hardware [2].

Additionally different attempts to improve interaction with the products assembly using virtual and augmented reality has been presented [15,53].

However, because of the current limitations of the enabling technologies, many problems have been reported from the experiences of the development and use of AR in industry. These problems are related with nonoptimal technology in fields such hardware, computer vision (segmentation, object recognition and rendering), tracking and registration, computer interaction and the adaptability to unexpected situations [20,33,39,54].

The combined issues in the enabling technologies make an incomplete and very limited technology that is unable to succeed in an industrial commercial world [13].

Therefore, in order to diminish the current limitations of technology such as cumbersome devices and avoid overload and over-reliance that are present in visual AR systems [54], the authors propose a method that augments the tactile perception to guide the assembly process.

The study performed in this research is aimed to support technologies of transfer learning on real-time, avoiding over-reliance on the technology and another well-known issue of continuous usage of AR devices.

1.2 Interactive design

In this article, we consider the interactive design and simulation like the one that allows bi-directional communication between the information of design and the real-life user. Where the information of the design is usually expressed in terms of computational models.

Hence most of the engineering activities are organized around virtual representations of products and processes. In which knowledge expert systems link classical Computer-Aided Design (CAD) systems with skills and know-how and are intended to spread the sharing of knowledge [12].

Therefore, in a further step, the interactive design will allow to directly connect the final users with virtual models allowing to easily adapt to new situations without incurring in more development time or new resources.

Thus, the proposed system works based on a pre-defined 3D models of the assembly and their visual appearance on real life is learned automatically by the system. Further, by itself (the system) is able to communicate in real-time the assembly of the object adapting to incorrect user's movements.

Where our approach is proposed to only require to set as input the 3D model of any 2D-like assembly to accomplish this task. In this way, the user could assembly any object without previous knowledge.

The guidance method was validated with a simplified assembly operation that is simulated with the task of putting together a puzzle. In the next section, previously developed

AR systems for assembly are reviewed, followed by the guidance method description (Sect. 3) and the development of the hapticAR application in Sect. 4. The validation test is presented in Sect. 5 and obtained results are discussed in Sect. 6. Finally, some conclusions are presented in Sect. 7.

2 Previous work

2.1 AR in assembly

A general overview of the current status of AR applications for assisting assembly operations was presented by Wang et al. [56] in which applications in three areas were reviewed: guidance, training, and simulation. Regarding assembly guidance it was found that although it is a classic problem existing for a while, there are still areas of improvement regarding instruction's interaction, context and user awareness in complex situations.

Some recent attempts of AR visual assembly guidance have been proposed. Anderson and Okamoto [34,35] proposed an AR system for guiding, through visual signals, the assembly of a pentomino puzzle, using a fixed tablet computer, between the user and the parts, like a see-through display and a processing unit. It uses markerless recognition and guides parts to a fixed position to finally verify the final assembly. This shows possible improvements in enabling technologies such as segmentation and real-time computation.

Similarly, puzzles have been largely used in AR for testing assembly implementations [25,34,40–42,44,46]. Some of the reasons for the use of puzzles simulating an assembly operation are that they allow to have control over the characteristics of the assembly and modify some tangible variables such as materials, dimensions, assembly path and also allows to define more "intangible" characteristics, for example, the assembly complexity or control the learning factor.

Besides a "well" defined design for assembly, AR has proven to be an effective solution for supporting assembly tasks. It was found that compared with printed instruction, AR reduced the error rate (82%) and time [48]. Also, it can reduce the mental workload on the assembly task by means of reducing head and eye movement between the assembly objects and the instruction, conducting to less attention switching. In addition, its use can reduce the mental activity on the transformation of object location and the amount of the information that the user needs [3,19,36,48].

Alternatively, main problems regarding technical issues were calibration, tracking, object recognition, portability, among others [11,60]. Regarding user interaction, over-reliance and overwhelming the user with virtual clues can cause the user's distraction [48]. In some tests, the users expressed that the Head Mounted Display (HMD) was

uncomfortable and had poor image contrast [3]. Those issues lead to some disadvantages such as the stress produced by a long-term usage of AR technologies [51].

In this way, a successful AR mobile system is the one that "enables the user to focus on the application or system rather than on the computing devices" [8], where the main nontechnical challenges are regarding the social acceptance. Here the goal is to not disrupt the user by being subtle, discrete and unobtrusive and this can be achieved by using a natural interaction and been fashion acceptable [8].

Some other areas of interest regard intelligent systems where the main challenge of assembly assisted by AR is to determine what, where and when the information will be displayed including the understanding of the surrounding assembly scene [36].

Ahmaniemi and Lantz [1] evaluated the task of target founding by scanning an AR environment with two conditions, pointing direction and proximity clues, finding that the performance between conditions does not differ much and a higher influence in the width and the distance of the targets (index of task difficulty).

Finally, Palmarini et al. [38] performed a systematic review of the current state of the art of AR in maintenance. Their main conclusions were that AR is not yet in a mature phase, where main problems are related to reliance, performance, and comfortable devices.

2.2 Vibrotactil guidance

On the other hand, vibrotactile feedback has been also proposed as support in AR in order to achieve sub-skills transfer for maintenance and assembly training, showing potential for skills training enhancement [57].

Similarly, Oron-Gilad et al. [37] evaluated the use of vibrotactile clues to guide an operator toward a target and examined the nature of the clues and their effect on target acquisition. They found that it is possible to reduce visual search by varying the pulse rate as a function of the target.

Further, Tan et al. [47] used a haptic back display using a 3-by-3 tractor array to send attention and directional information to the user, decreasing the reaction time by 41% and increasing average reaction time by 19%. This becomes useful for users needing to pay attention to the smallest areas in large and complex visual displays.

A visual search task in multiple screens was evaluated with auditory, haptic, combinatory and no clues. Results showed a significant and substantive improvement in performance with combinatory clueing. Tactile clueing leads to increasing speed performance [18].

Also, other attempts have been proposed for hand position and orientation guidance. Sergi et al. [43] used four vibrotactile stimulators to provide user feedback about the user's forearm direction in the Cartesian space. Showing to be a valuable tool for motion guidance and compared to visual guidance alone, vibrotactile plus visual guidance improved the positioning accuracy.

A 3D Orientation Guide (3DOG) was proposed by Guo et al. [17] composed by 3 vibrating motors in order to communicate forearm postures and inform correct position. This provided an effective and intuitive feedback, being promising for its wearability, effectiveness and low cost.

A study comparing verbal versus tactile information to guide the hand to a target position was developed by Lepelley et al. [30] showing that spatial information can be conveyed using the tactile channel (93.27% of participants identified the directions) and no significant differences were found between the two channels in terms of number of precision or participants kinematics.

Additionally, vibrotactile also has been used as a support of navigation through unfamiliar places. Straub et al. [45] proposed a tracking hardware and a vibrotactile waist belt for an indoor navigation system. Results showed a precise walking by using vibrotactile clues alone. Similarly, Bosman et al. [4] developed GentleGuide, a system for delivering haptic guidance in the wrists for pedestrian guidance. This showed a promising technology in terms of reducing disruptiveness of technology, which is a key aspect in the technology acceptance in AR.

Also, a vibrotactile feedback was used to teach how to balance a simulated inverted pendulum where the haptic information encodes the state and was different from the visual. This showed an improvement by three times in the time to failure and even the effect persisted after the feedback was removed [52].

A Vibrotac device was used to guide the translation and rotation of virtual objects and vibrotactile and verbal guidance were compared without visual information. The results showed that although the users were faster using the verbal guidance, the movement accuracy was similar, probably due to more time is required to interpret the vibration signals for translation. The opposite occurred with rotational information where the user reached the target angle faster and with more accuracy using vibrotactile information [58].

Further, a dynamic tactile clueing system, proposed by Lehtinen et al. [29], couples the hand position with the scene position in order to guide actively the hand towards a target using tactile feedback. Substantial improvements were found compared to visual search only and demonstrated that the effect of visual complexity of the scene can be eliminated.

Similarly, Nii Mante et al. performed a study to evaluate the performance of blind individuals to grasp objects using vibrotactile clues controlled by a real-time system using computer vision. In their experiments, their results were compared against verbal guidance. As result, the vibrotactile guidance led to better performance than the auditive guid-

ance. The further recommendation followed in this research was to use a confirmation feedback [32].

Jose V. Salazar Luces et al. proposed a guidance system based on a Phantom-Sensation paradigm. Their system uses an elastic band with six vibration motors to guide the wrist rotation. In their results, their shown a successful guidance in any user position creating virtual stimulus between the equally spaced motors [31].

AR systems for supporting assembly tasks have shown some benefits such as error, movements and mental stress reduction. And, although the enabling technologies that compose AR systems have been improved over the last years, there are still many unsolved issues such as portability, resolution, object recognition among others that cause long-term use problems and prevent their use outside of the laboratories.

On the other hand, vibrotactile guidance has shown promising results that are useful in the assembly guide and complements some aspect of AR visual systems. Some of them are low disruptiveness, visual complexity reduction, portability, and minimizing user's dependability to the technology.

Thus, the next section presents a vibrotactile system (HapticAR) for the guidance of the assembly of a 2D puzzle, aiming to minimize some of the current enabling technologies issues by taking the advantage of haptic guidance.

3 Vibrotactil guidance method

The guidance method proposed in this research was inspired by the Gestalt psychology that stands that when a number of stimuli are presented we do not experience them as individual things but instead as a larger whole [59]. Where in its origins, Max Wertheimer discovered the phi motion, where a series of still images presented rapidly in a sequence could be perceived as continuous movement [55].

Therefore, our guidance method proposes a series of individual stimuli (vibration clues) that aims to be interpreted as a global instruction for positioning the parts of the assembly in the correct place.

The total assembly of the element is accomplished by positioning each one of the parts in the correct place with the correct rotation. Thus, we propose to divide the final positioning of each part in three activities: (i) select the correct part, (ii) rotate in the correct angle and (iii) place in the correct part.

Part selecting The user starts to scan the parts using the hand with a pointing hand gesture. When the user pointing finger passes over the desired part to be assembled the system gives a vibration clue.

Rotation For the rotation two types of clues (vibrational) were used. The first one to give a feedback on the angle of rotation to simulate the rotational feedback of physical knobs. And finally, when the user achieves the correct rotation a different vibration pattern was applied. The visual representation is shown in Fig. 9 were each circle represents a haptic clue.

Position For the position, first is indicated the position in the x direction and later the y direction to the final position. Similarly, the user feels a feedback each defined space if he or she is directing in the right position if not no feedback is returned. If the user takes a wrong path, the instruction is recalculated. Finally, when the user arrives at the correct direction a different haptic clue is applied. The visual representation of the positional instructions are shown in Fig. 7.

## 4 Technical system description

The details of the implementation of HapticAR are presented in this section. The proposed system works in two stages. An offline stage where a model for performing object classification is trained, and an online stage, where the system guides the user in the assembly of the puzzle.

In the offline stage, a synthetic dataset is built using the 3D models of the parts and generating renders under different points of view. For each render of the part, the Hu moments of each sample is calculated and stored as the features of the parts. Later with the dataset, a logistic regression model is trained. The model is trained to recognize each part of the puzzle given as input its Hu moments.

Additionally, for each part, the centroid, a representative vector and the relative position of each part regarding the previous are calculated and stored. Likewise, in order for the system to separate the background from the parts and the user's hands, in the offline stage, the mean and standard deviation of the background and the parts puzzle parts are learned.

Later in the online stage, the system receives as input the frames of the camera. For each frame, it separates the pixels corresponding to the background, parts and user hands. For each part of the puzzle, the Hu moments are calculated and recognize each one predicting the part using the trained model.

Afterward, it verifies that all the puzzle parts are in the visible area and start guiding the user by transmitting different vibration patterns when the user passes over the target
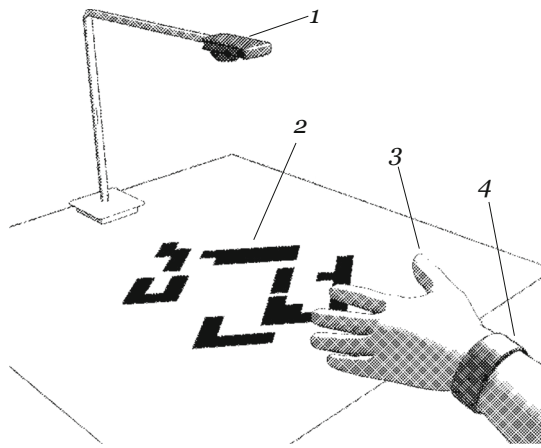
Fig. 1 HapticAR general layout. Components: (1) Webcamera, (2) Puzzle parts, (3) User hand, (4) Vibrotactile wristband

part. Once the user's hand is over the right part, the system gives vibration clues to guide the user to drag of the part to the correct position.

Finally, the system verifies the position of the part or the performed assembly until reach a correct solution.

4.1 Hardware configuration

The main layout considers a web camera (1) located on top of the assembly table, a flat puzzle (2) and a vibrotactile device (4) placed on the wrist of the user (3), as it is shown in Fig. 1. The system uses a web camera (1) as an input system to recognize the parts and the user hands (3). This information is processed and, according to the assembly situation, the system sends an electrical signal, using the audio port, that is amplified and converted to vibration using electrical motors placed in the user's wrist (4).

The acquisition system (No. 1 in Fig. 1) is a Logitech®Webcam C905, which has a Carl Zeiss®Optics with autofocus, a native 2MP sensor that allows having up to 1600×1200 MP and a frame rate of 30 FPS. Besides, the camera is positioned parallel (and fixed) to the assembly surface in order to capture the top view of the parts and the user's hands.

All the computations are performed in a laptop connected to the webcam and to the vibrotactile device. The laptop has an Intel®Core i5-2410M 2.30 GHz processor with 4GB of RAM and an integrated video card.

The vibrotactile device is a wristband (Fig. 2) with two vibration motors attached that are connected to the sound port of the laptop. The vibration motors are a standard coin-like motor of 10 mm of diameter by 4 mm of depth with a working voltage of 2.5–4.0 V and they are usually found in mobile phones.

In order to power the vibrator, the amplifier circuit presented in Fig. 3 was used with an input of 9V battery, a
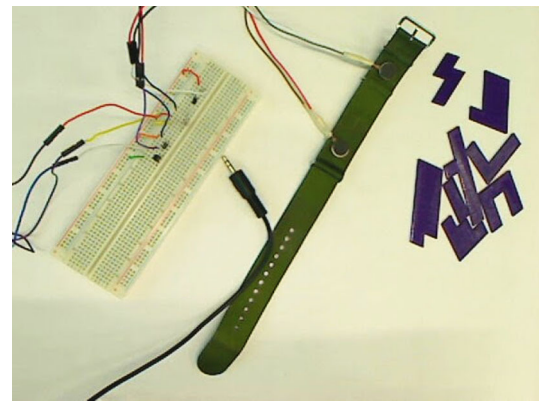
Fig. 2 Vibrator device, amplifier circuit and assembly parts
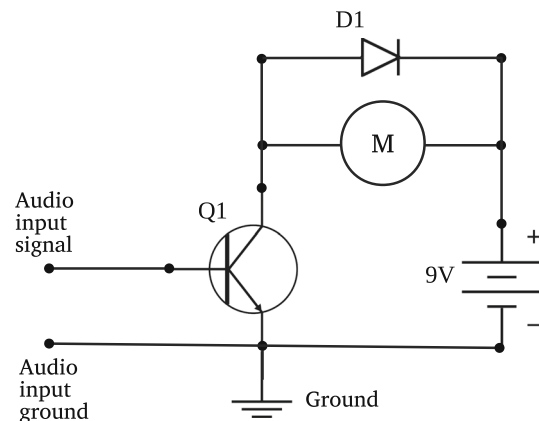
Fig. 3 Vibrator wrist amplifier electronic scheme

transistor (Q1) and a diode (D1). Each different motor is controlled by a sound channel (left or right) of the audio port.

This configuration allows having the flexibility of trying different vibration patterns and, at the same time, keep similitude with commercial devices as smartwatches.

Additionally, two different vibration patterns, that are transmitted by the audio port of the computer, were used, as it is shown in Fig. 4: one of 10 Hz of frequency (Lower wave in Fig. 4) for the clues and the other one of 5 Hz of frequency (Top wave in Fig. 4) for communicating that the activity is performed correctly. Both patterns have a square waveform and are played with a duration of 100 ms and 500 ms at 44.1 kHz, respectively.

The user performs the task while is sitting, with a constant lightning and is allowed to use both hands to perform the assembly.

4.2 Segmentation

Segmentation is the process of defining parts of the image with internal similitude, such as texture or statistical simi-
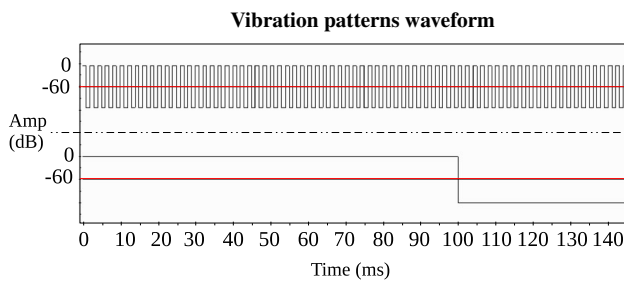
Fig. 4 Superior part of the image: vibration wave form used for correct action. Lower part of the image: clues wave form

larity [27]. In our case, 3 elements need to be isolated: the background, the parts, and the user's hands.

The segmentation is the base of the system since it defines the input to the subsequent process [34]. Thus, in order to avoid interference of lightning or background changes, it is used an average backgrounded method [5] which learns both the average and an approximation of the standard deviation of each pixel as a background model.

This method works for a static camera and steady background with slight changes in lighting (indoor) and allows to have a robust segmentation process.

The segmentation is made in two processes, a learning, and a thresholding process. In the learning phase, in a set of frame-time n, two matrices containing the information of each pixel regarding the pixel intensity mean $\bar{I}$ and standard deviation (Eq. 1) is approximated as the frame to frame absolute difference:

$$\sigma \approx \frac{\sum_i^n |I_i - I_{i-1}|}{n} \tag{1}$$

Afterward, low and high threshold values for each pixel are defined (Eq. 2):

$$T = s\sigma \pm \bar{I} \tag{2}$$

Thus, after segmentation, the background and foreground elements are defined. In order to differentiate the hand and the parts, two similar processes are performed. The hands' segmentation is achieved using a skin color segmentation that is described in Sect. 4.4.

Assuming that we already know what parts of the image are the background and user skin, the part color is also learned. First, a recognition of the parts is made and after identifying a minimum number of parts, the color of the parts is learned using the same approach used for the background learning.

4.3 Object recognition

The proposed system uses a markerless approach in which the parts are described by features and learned in an off-line process.

Object Recognition (OR) can be defined as the act of evaluating the presence and location of objects based on some knowledge about their appearance [50]. An OR system usually has the next following components [23]:

1. Database: where the model of the objects is contained and information about their characteristics (features) that are important for recognizing them
2. Feature detector: functions applied to the images to extract the features
3. Hypothesizer: Reduce the candidates of possible objects
4. Hypothesis verifier: The system defines the object by the highest likelihood regarding database information

One of the most important aspects is to define how to represent the objects. It was decided to represent the parts that compose the puzzle using the Hu moments as features [21]. Moments are scalar quantities that can describe the shape of a probability density function, where an image can be seen as a piece-wise continuous real function $f(x, y)$ defined in $D \subset \mathbb{R}^2$, and can be defined (Eq. 3) the general moment M_{pq} of an image $f(x, y)$ of order $(p + q)$ [14]:

$$M_{pq} = \iint_D p_{pq}(x, y) f(x, y) dx dy \tag{3}$$

Thus, the raw moment M_{ij} of a binary image $I(x, y)$ can defined as (Eq. 4):

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \tag{4}$$

The moment about the variable mean is call central moment and it is where the centroid (Eq. 5) is in an image:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}} \tag{5}$$

The central moments (Eq. 6) are defined as:

$$u_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{6}$$

Where the normalized moments (Eq. 7) can be computed as:
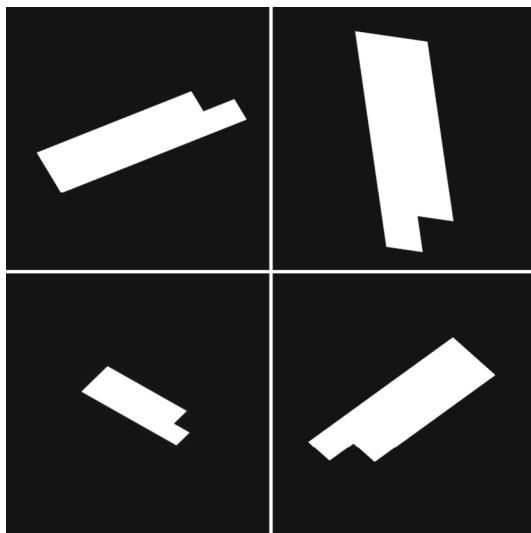
$$n_{ji} = \frac{u_{ji}}{u_{00}^{1+((i+j)/2)}} \tag{7}$$

Fig. 5 Example of transformation used in the synthetic image training

Fig. 6 Normalized image moment h_1 of the different parts of the puzzle under scale and rotation transformation versus the side resolution l of the image (square)

In 1962, Hu proposed seven moments (Eq. 8) invariant to Translation, Rotation and Scale (TRS) based on the work of Boole, Cayley and Sylvester [21,49] on the theory of algebraic forms defined as:

$$
\begin{aligned}
h_1 &= \eta_{20} + \eta_{02} \\
h_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
h_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
h_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
h_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
&\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
h_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12}) \\
&\quad (\eta_{21} + \eta_{03}) \\
h_7 &= (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] - \\
&\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}
\tag{8}
$$

The moment invariants have become the most important and frequent used shape descriptors. Despite that, they can just be used as global descriptor preventing their direct use for occluded objects [14]. Therefore, each part of the puzzle is represented by a seven array of Hu moments.

A synthetic training set m was built simulating the associated transformation of the puzzle parts and calculating their respective moments x as can be seen in Fig. 5. A total training set containing 144,000 images was used.

One of the major issues to be considered about the use of Hu moment as features is that they are invariant for continuous functions, but in discrete images, moments change over geometric transformations (scale and rotation). Those fluctuations decrease when the resolution of the image increases until reaching a threshold but at the same time increasing computation [22].

As can be seen in Fig. 6, where the different puzzle parts were scaled 50 times and in each scale were rotated in 360°
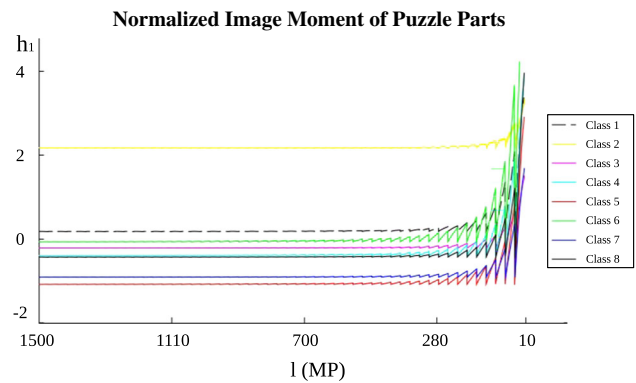
with an initial resolution of 1500×1500 MP and a final resolution of 10×10 MP, the small fluctuation starts to be more noticeable when the resolution decreases. To overcome this issue a minimum resolution of 280×280 MP was used.

Additionally, a logistic regression model is used to define a classifier, where the estimated probability (Eq. 9) of be a part h_θ of a new set of normalized moments x, given a set of learned parameters θ:

$$
h_\theta(x) = \frac{1}{1 - e^{-\theta^t x}}
\tag{9}
$$

The cost $J(\theta)$ of the prediction h_θ against the real part y, using a given set of parameters is estimated by (Eq. 10):

$$
\begin{aligned}
J(\theta) = -\frac{1}{m} \Bigg[&\sum_{i=1}^{m} y^{(i)} log(h_\theta(x^{(i)})) \\
&+ (1 - y^{(i)}) log(1 - h_\theta(x^{(i)})) \Bigg]
\end{aligned}
\tag{10}
$$

Finally, an optimization process is done to minimize $J(\theta)$. These learned parameters (θ and normalization parameters) are used later in the recognition phase using Eq. (9) to predict the probability for a region to be a part. Then, a threshold process is followed to consider only the ones with high probability.

4.4 User recognition

The user is supposed to perform the assembly using both naked hands. Thus, the first step was to identify the pixels belonging to the user. A skin segmentation is performed by a method proposed by Kovac et al. [26] in which they use the 3D color space (RGB) to detect skin color pixels. By using heuristics rules (Eqs. 11–14) to classify each image pixel and its mean to work better under standard daylight illumination

Fig. 7 Hand segmentation and gestures performed by the user for the part positioning to target point through active clues. The user selection point is the tip of the fingers and clues are activates one the finger pass over them
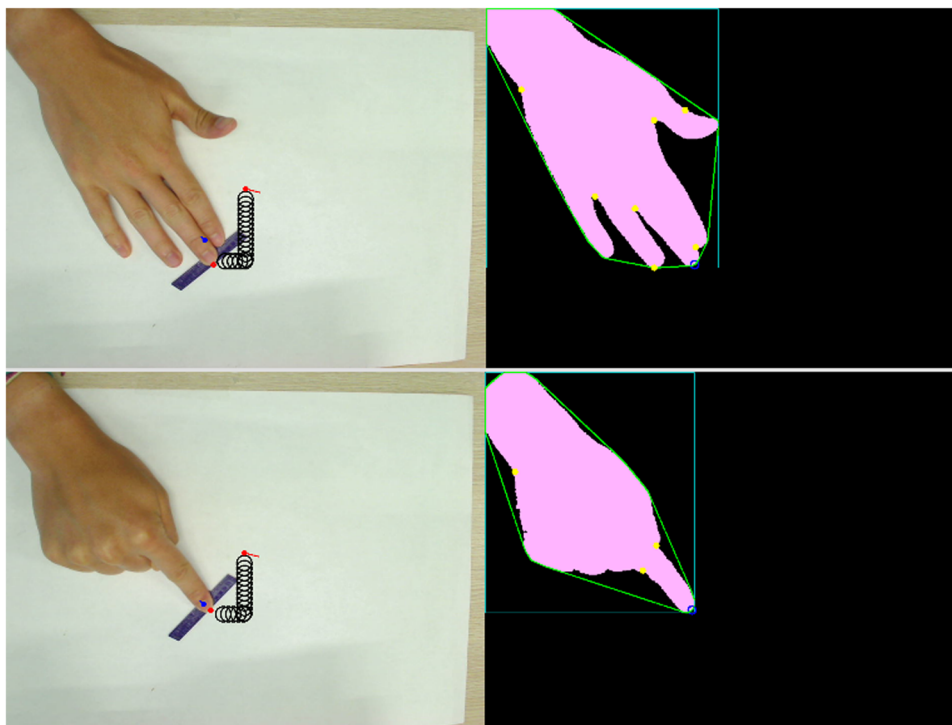
(CIE illuminant D65). The rules are defined as follows for each pixel in the $[RGB]$ plane:

$$R > 95 \quad \& \quad G > 40 \quad \& \quad B > 20 \tag{11}$$

$$max(R, G, B) - min(R, G, B) > 15 \tag{12}$$

$$|R - G| > 15 \tag{13}$$

$$R > G \quad \& \quad R > B \tag{14}$$

The selection and manipulation of the puzzle parts can be done by two different gestures, pointing and extended hand, as can be seen in Fig. 7. A selection point is defined in the tip of the fingers and is the one that is going to establish what part the user is pointing and moving.

In order to define the selection point, the convex hull which is the smallest convex region that encloses the hand pixels is calculated [16] and it is approximated to a polygon [10].

After that, all the angles of the polygon formed by the convex hull are computed. The candidates for a possible selection point are selected by an angle range threshold and finally, a point is designated by removing borders candidates.

4.5 Parts position and orientation

For each part of the puzzle, the application defines its relative target position and rotation accordingly, based on an input image of the desired final puzzle assembled. In this case, we define the assembly configuration showed in Fig. 10, but any other configuration could be selected.
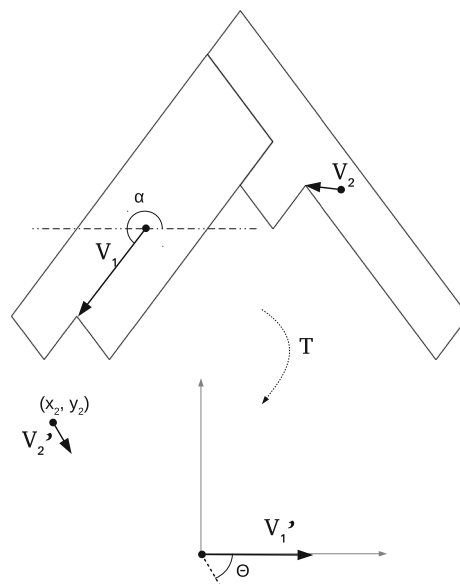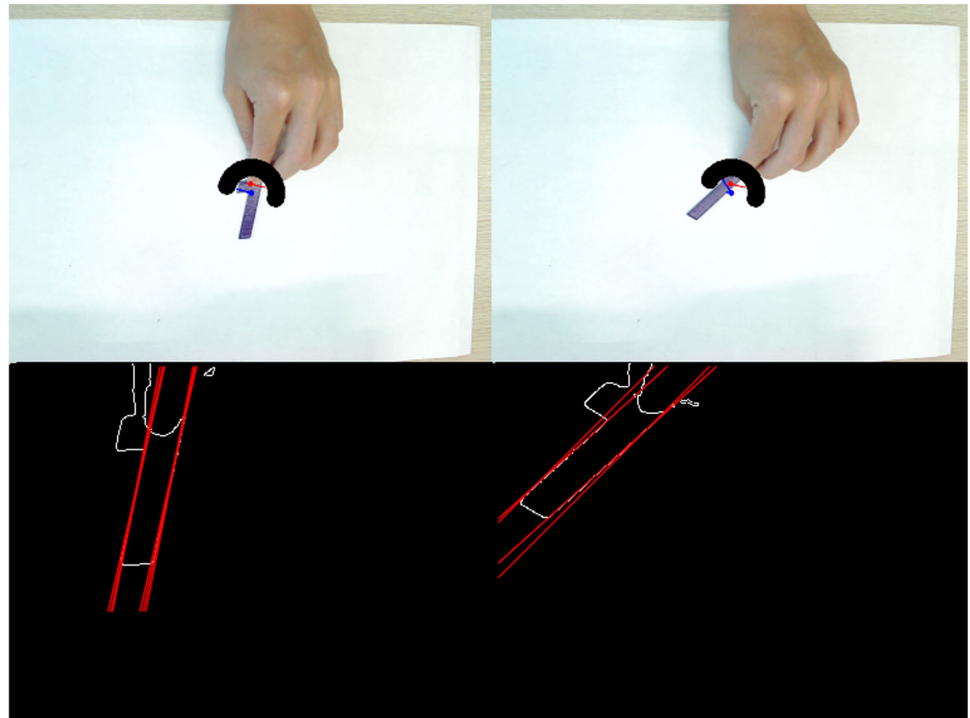

Fig. 8 Puzzle assembly learning of part 1 and 2. Parts represented by vectors and relative coordinates and angles relative to previous part

The assembly learning is performed when the application starts by following the next steps. Where each puzzle part is represented by a vector. The vector tail is defined by the part centroid (Eq. 5) and the tip by the nearest point of the polygonal part representation. The relative position of the consecutive part (v_2) regarding the previous one (v_1) is calculated by first applying the transformation T (Eq. 15) that aligns the v_1 with the x-axis, as it is illustrated in Fig. 8.

Fig. 9 Rotation guidance to fixed target angle using virtual path of vibrational clues and real time angle

$$T = \begin{bmatrix} cos(\alpha) & sin(\alpha) & -v_{1x}cos(\alpha) - v_{1y}sin(\alpha) \\ -sin(\alpha) & cos(\alpha) & +v_{1x}sin(\alpha) - v_{1y}cos(\alpha) \\ 0 & 0 & 1 \end{bmatrix} \quad (15)$$

For the next parts, the sum of the previously assembled regions will be taken as one new part since it is not possible to identify individual parts once they are joined together.

This configuration allows the user to assemble the puzzle in any location and rotation, as long as starts from the first part. Additionally, this allows us the flexibility of defining any puzzle configuration.

Finally, in order to get an approximated angle of the rotating part and overcome the partial occlusion generated by the hand, a very useful technique is the Hough transform [6]. In which a window surrounding the part position is settled and the edges are calculated using canny edge detector [7] only on the puzzle parts region. Then, a voting system for all the possible line parameters, in polar coordinates (ρ, θ) of each edge pixel, is performed, and the parameters with more votes are selected.

Once the lines are identified, given the characteristics of the parts, only parallel lines are used for calculating the average change of angle between frames. This change is subtracted to the angle of the vector that represents the part.

Once the target position is reached, the rotation clues are set similarly to the position clues. A circular path of circles is set every degree from current vector part rotation until target rotation, as it is shown in Fig. 9. Further, a real-time part rotation is estimated to provide feedback while the user rotates the part.

4.6 Model updating

Finally, one of the properties of the proposed system is that is flexible for new objects, parts and assembly routes. The input of the system is the 3D model of the assembly and the index of each one of the parts equivalent to the order of assembly.

The process for training for a new assembly is as follows:

1. Synthetic data generation: Using the 3D model generate a dataset of the parts based on affine and perspective transformations of the 2D projection of the parts.
2. Train the object detector using the previous dataset as presented in Sect. 4.3. This model predicts the probability that a blob is a part of the assembly. As the predictor return the index of the part, this information is used for the order of assembly as well.
3. Train the relative position and orientation of each one of the parts, this process is performed online as it takes virtually no time, and consist of storing the relative position between the parts.
4. Learn background, objects color and lighting. This step is also performed online at every time the application starts in order to adapt to changing conditions (Sect. 4.2).

Based on this configuration, the system allows the changing of (i) order of assembly by changing the index of the parts, and is not required to re-train the prediction model, (ii) change the assembly form by reordering the 3D assembly, this also is performed without re-training, (iii) visual appear-
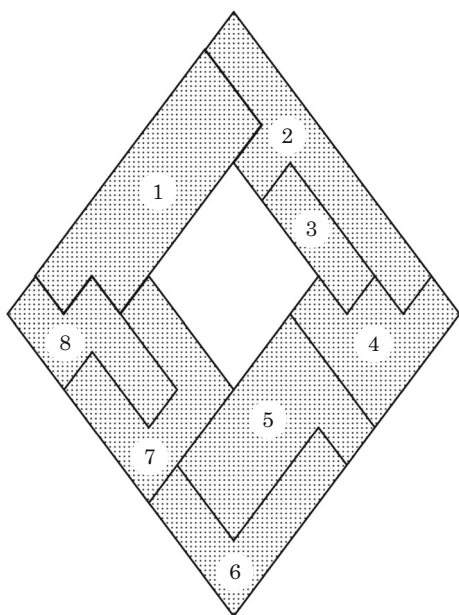
Fig. 10 Assembled puzzle

ance such as background, lighting and parts color and (iv) assembly parts: by retraining the object predictor.

5 Experiments

In the next section, the different tests performed to validate the functionalities and interaction of the proposed application are presented. In order to verify the different functions of the application, two types of test were performed, a system functions' test and a user interaction test.

In order to evaluate the proposed system, we have defined a "dummy" assembly operation where a user has to solve a tiling puzzle of an unknown form, as it is shown in Fig. 10. The parts are spread over the table without overlapping each other. The wristband and the 3D printed parts used for the test are shown in Fig. 2.

5.1 System functions test

The first test performed was the object recognition where all the different parts of the assembly were randomly arranged and the system identified them. A total of 43 different random arrangements were performed, considering the 8 parts of the puzzle.

Further, the parts were placed avoiding occlusion or joined parts. Also, no parts different to the ones of the puzzle were added. The lighting of the test was controlled in an indoor environment and a clear background, as it is shown in Fig. 11.

The second functional test is intended to evaluate the verification of correct assemblies of the system. For each step of the assembly, six different assembly configurations were made, in which three of them were correct assemblies but in different locations and rotations and other three were incorrect assemblies, as it is shown in Fig. 14. A total of 42 assemblies were tested with the same environment, lighting, and background that the previous test (Fig. 12).

5.2 User interaction and performance test

Interaction tests involved a user performing different activities required to complete an assembly operation. Three main activities were defined: (1) identify a part from the puzzle, (2) rotate the part to a given angle and (3) move the part to a target position in a (x, y) plane. Additionally, the vibrational clues were the only available information to perform the tasks and no other additional information, such as visual or auditory, was used. Also, the user was not deprived of the sight and (s)he was allowed to see the parts in the test.

A total of 39 users, with an age ranging from 18 to 36, and an average age of 24, participated in the test. Most of them were Engineering students, 33% of them were women, and only one of them reported to be left-handed. Finally, none of them was familiarized or used this kind of technologies before.

Before starting the test, the two different vibration patterns were applied to the user using the vibrotactile device and they were asked if they were able to differentiate them. Additionally, none of the users had previous knowledge or had ever tried the application before.

In the first part of the test, all the parts of the assembly were spread randomly and the user was asked to pass the hand over the parts with a pointing gesture (equal to showed in Fig. 7 and the parts in a similar layout showed in Fig. 11) and the task consisted in that the user has to tell what was the part that the system was indicating. When the user's pointing finger was placed near the center of the target part, the systems transmitted a vibration clue to the user.

For the second and third part of the user interaction test, the user is supposed to position one part in a target position $(x, y, angle)$. The target position is at $angle = 15°$ of rotation regarding the vector part and the x axis, considered correct with a $\pm 10°$ range. The correct (x, y) position is located in (480, 360) px from the top, left corner of the camera view.

Therefore, the user is first asked to rotate a specific part following the clues provided by the device and to indicate when (s)he thinks the part is in the right rotation, as it is shown in Fig. 9. For the limitation of the camera as the only input system, the user is asked to rotate the part from any corner to avoid full occlusion. Additionally, it is suggested to the user if in any moment (s)he does not feel the vibration for a long period of rotation, to remove the hand from the table and start again to rotate the part.

Fig. 11 Randomly placed parts for system OR test and initial layout for user part selection test: without occlusions or joined parts in a clear back ground with constant lighting
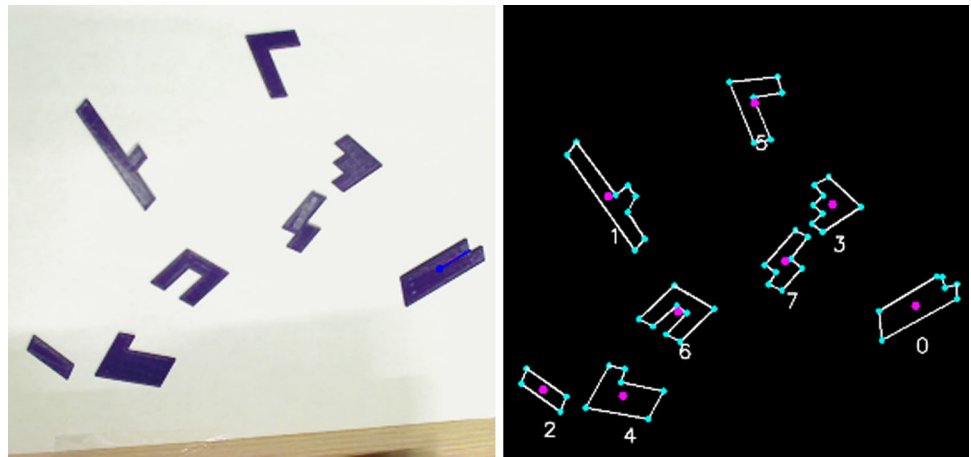
Fig. 12 System verification test of the 4th step of the assembly of the puzzle

(a) *Incorrect Assembly*   **(b)** *Correct Assembly*
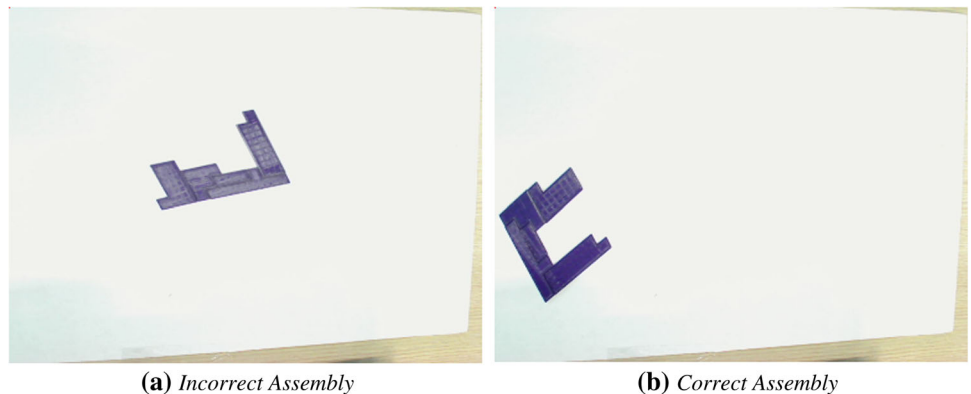
Similarly for the last test, after the user sets the rotation, (s)he is asked to drag the part into the target position (x, y) as it is shown in Fig. 7 with a random initial position. Additionally, the users were informed that the clues will be the first horizontal, along with the x axis and then vertical. And, to stop when they feel the vibration pattern for the correct position.

6 Results and discussion

The proposed Object Recognition of the system was tested with a 2D puzzle, of 8 parts in 43 random arranges without occlusion. A total of 344 parts were required for recognition of the system. The 95% of the parts were recognized and 79% of the arranges were identified. Most of the recognition problems were due to that the parts were positioned at the corners of the video or to the glossiness of the parts. Such causes are expected for a system when global features are used for OR, that is not robust against occlusions or specularities.

Additionally, for the assembly verification system, the 7 steps required to perform the assembly (Fig. 10) were tested with 3 correct and 3 incorrect assemblies (Fig. 14). A total of 21 assemblies were tested, in which all the correct assemblies

were identified and 6 incorrect assemblies were confused, obtaining a 78% of recall.

Evaluating the user interaction, all the users at the beginning of the tests were able to identify the two different patterns correctly. In the first test, they were all available to select the correct part indicated by the system. The only question was about the pointing gesture (Fig. 7) since it was the first time for them using the system, some of them presented some confusion about how to do it.

Regarding the positioning tests, two main issues were identified. The first one was related to the confusion of the vibration pattern for corrective actions when the user had not reached the target position with the clue vibration. This occurred when the user moved too fast and several vibration clues were applied. These occurrences, 2 in rotation and 6 in location, were treated as outliers and removed from the analysis.

The second issue was about the target rotation angle, where the system guided the user to a totally opposite angle 195°. This occurs because of the way the representative part vector is calculated, the geometry of the part used for the experiment and the low camera resolution. The system calculated a vector regarding the mass center and the nearest vertex (presented in Sect. 4.5). Thus, in these cases, the target rotation was changed to 195° since the aim of the experiments

Fig. 13 Flipped representative vector part due to camera resolution and environmental conditions
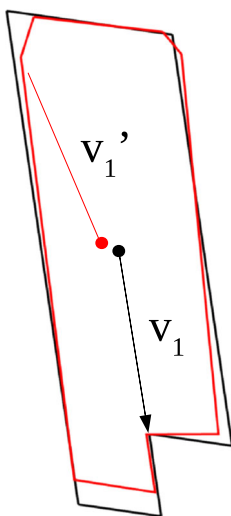
was to evaluate the interface and the possibility to guide a user.

In parts near to be symmetric, especially small parts, it is easy to recover the shape, with some deformation, given the low camera input resolution. Therefore, the center of mass could have some offset and additional vertexes could appear and be used for defining the representative vector, as it is shown in Fig. 13.

Further, in the second user interaction test, rotate the part to a target angle, the 36% of users stopped at the alert range with an average angle of 30.9° and an average time of 49s. Also, the distribution of the normalized error (difference between target rotation and user final rotation) is shown in Fig. 14a, with a sample size of 37, with $\sigma = 0.15$, $\mu = -0.98$ and a Shapiro–Wilk normality test p value of 0.096, with a significance level of 0.05.

Most of the rotation errors were due to totally or partially part occlusion while rotating. Therefore, the system was not able to recover the real part rotation and there was an offset between the real rotation and the system rotation. This gap of rotation is fixed when the user removes the hand of the part as long as (s)he has not reached the target, that is the case of the test. Therefore, this error could be minimized if more rotation iterations are considered instead of one iteration.

On the other hand, regarding the movement of the part to the correct position (x, y), the average Euclidean distance to the target was 44.5 px and it takes an average time of 84.2 s for the users to reach the target.

The positioning test was analyzed for each axis independently. Thus, for the position of the part in the x axis the error, by user, was calculated by $e_x = (t_x - u_x)/w_x$, in which t_x is the target position in the x axis, u_x is the position reached by the user in pixels, and, the error is normalized by the total available space, w_x, in this case, the width resolution of the web camera.

The distribution of the normalized error in x, with a sample size of 33, is shown in Fig. 14b, with $\sigma = 0.026$, $\mu =$

-0.006 and a Shapiro–Wilk normality test p value of 0.5, with a significance level of 0.05.

Similarly, for the y axis, the error distribution is shown in Fig. 14c, with $\sigma = 0.03$, $\mu = -0.009$ and a Shapiro–Wilk normality test p value of 0.6, with a significance level of 0.05.

Finally, some considerations about the user interaction tests were identified. One of them was related to the different skin tones, that is required to calibrate the white balance of the camera. Additionally, another cause of confusing the clues by the users was due to the use of nails polish, since for some user was confusing, or (s)he performed with large variations, the pointing gesture leading to the system to detect the pointing in the incorrect place. On the other hand, other users moved their hand too fast, while dragging the part, leading to confusing the vibration patterns. And, for some user was difficult to understand the change of axis while performing the dragging.
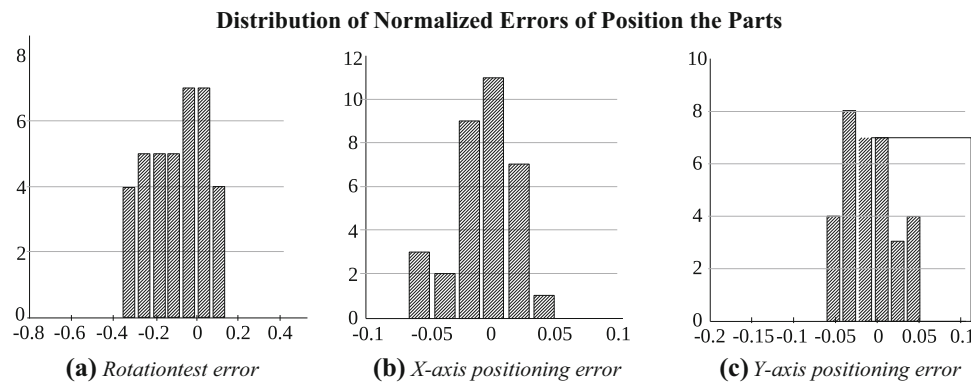
It was observed that the performance of movements and searching style variate largely from user to user, especially reaching the location (x, y). This depends on each individual deduction capacities with the current clueing system. But, in general, a good perception by the user, while using the system and after the test, was perceived.

More intelligent clues definition and more proactivity to user changes could be the way for improving the system. Also, extending its usage to 3D objects. Additionally, improvements in segmentation and object recognition are required in order to allow a more deeply study of how this guidance affects the learning and dependency of the system in the assembly process.

7 Conclusions

Here is presented an application, so-called HapticAR, combining both Augmented Reality and Haptic technologies, for assembly guidance, that aims to avoid some of the currents withdraws of the use of AR by augmenting another human sense. The main conclusions are:

- The AR systems are a mix of interconnected enabling technologies (Object Recognition, segmentation, hardware, user interaction), and the errors in one of them affects or produces errors on the other process. This is largely seen on how hardware and OR, that is the base of the proposed system, influenced the performance of the users on the tests.
- By augmenting another human sense, some of the current issues of visual AR are improved, such as cumbersome hardware, low resolution, the field of view and, additionally, another induced related issues, as fatigue and stress. Nevertheless, still, there is a high reliance on hardware

Distribution of Normalized Errors of Position the Parts

(a) *Rotationtest error*   **(b)** *X-axis positioning error*   **(c)** *Y-axis positioning error*

Fig. 14 **a** Distribution of the normalized rotation test error (difference between target rotation and user final rotation) $\sigma = 0.15$, $\mu = -0.98$ and a Shapiro–Wilk normality test p value of 0.096. **b** Distribution of the normalized x axis positioning test error $\sigma = 0.026$, $\mu = -0.006$ and a Shapiro–Wilk normality test p value of 0.5. **c** Distribution of the normalized y axis positioning test error $\sigma = 0.03$, $\mu = -0.009$ and a Shapiro–Wilk normality test p value of 0.6

to perform the OR that is largely affected by the camera resolution (Sect. 4.3).

- In contrast to auditory or visual communication, the information transmitted with vibration is more limited. Therefore, it is more evidence that the user is required to use more deductive capacities to reach the target position. Consequently, it could affect how the users are focused on the task and how they learn new skills.

- The proposed system was able to recognize 95% of the parts of the 344 used in the Object Recognition test. Most of the errors were due to parts being cropped by being out of the camera and by the specular reflections. In order to recognize these variations and other such as the ones generated by changes in the point of view, occlusions or reflections are required to recreate them in the synthetic training.

- We have developed a low-cost system that is able to guide the user during the assembly of flat elements by performing an object recognition (accuracy 95%) and generating interactive haptic clues to guide the user to reach a target position (mean error of 0.15 rotation and x-axis mean error of -0.006 and y-axis of -0.009) and to validate the assembly (78% recall). In this way, this system could guide the user by its own or it could be the support system for other Augmented Reality approaches.

References

1. Ahmaniemi, T.T., Lantz, V.T.: Augmented reality target finding based on tactile cues. In: Proceedings of the 2009 International Conference on Multimodal Interfaces, pp. 335–342. ACM (2009)
2. Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B.: Recent advances in augmented reality. IEEE Comput. Graph. Appl. **21**(6), 34–47 (2001)
3. Baird, K.M., Barfield, W.: Evaluating the effectiveness of augmented reality displays for a manual assembly task. Virtual Real. **4**(4), 250–259 (1999)
4. Bosman, S., Groenendaal, B., Findlater, J.-W., Visser, T., de Graaf, M., Markopoulos, P. : Gentleguide: an exploration of haptic output for indoors pedestrian guidance. In: International Conference on Mobile Human–Computer Interaction, pp. 358–362. Springer (2003)
5. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media, Newton (2008)
6. Hough, P.V.C.: Method and means for recognizing complex patterns. US Patent 3,069,654, 18 Dec (1962)
7. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. PAMI **8**(6), 679–698 (1986)
8. Carmigniani, J., Furht, B.: Augmented reality: an overview. In: Handbook of Augmented Reality, pp. 3–46. Springer, New York (2011)
9. Caudell, T.P., Mizell, D.W.: Augmented reality: an application of heads-up display technology to manual manufacturing processes. In: Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, 1992, vol. 2, pp. 659–669. IEEE (1992)
10. Douglas, D.H., Peucker, T.K.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartogr. Int. J. Geogr. Inf. Geovis. **10**(2), 112–122 (1973)
11. Elvins, T.T.: Augmented reality: "the future's so bright, i gotta wear (see-through) shades". ACM SIGGRAPH Comput. Graph. **32**(1), 11–13 (1998)
12. Fischer, X., Coutellier, D.: The interaction: a new way of designing. In: Research in Interactive Design, pp. 1–15. Springer, Paris (2006)
13. Fite-Georgel, P.: Is there a reality in industrial augmented reality? In: 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 201–210. IEEE (2011)
14. Flusser, J., Zitova, B., Suk, T.: Moments and Moment Invariants in Pattern Recognition. Wiley, Hoboken (2009)
15. Gonzalez-Badillo, G., Medellin-Castillo, H.I., Lim, T.: Development of a haptic virtual reality system for assembly planning and evaluation. Procedia Technol. **7**, 265–272 (2013)
16. Graham, R.L., Yao, F.F.: Finding the convex hull of a simple polygon. J. Algorithms **4**(4), 324–331 (1983)
17. Guo, W., Ni, W., Chen, I.-M., Ding, Z.Q., Yeo, S.H.: Intuitive vibro-tactile feedback for human body movement guidance. In: 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 135–140. IEEE (2009)

18. Hancock, P.A., Mercado, J.E., Merlo, J., Van Erp, J.B.F.: Improving target detection in visual search through the augmenting multi-sensory cues. Ergonomics **56**(5), 729–738 (2013)

19. Henderson, S.J., Feiner, S.K: Augmented reality in the psychomotor phase of a procedural task. In: 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 191–200. IEEE (2011)

20. Hincapié, M., Caponio, A., Rios, H., Mendívil, E.G.: An introduction to augmented reality with applications in aeronautical maintenance. In: 2011 13th International Conference on Transparent Optical Networks (ICTON), pp. 1–4. IEEE (2011)

21. Ming-Kuei, H.: Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory **8**(2), 179–187 (1962)

22. Huang, Z., Leng, J.: Analysis of Hu's moment invariants on image scaling and rotation. In: 2010 2nd International Conference on Computer Engineering and Technology (ICCET), vol. 7, pp. V7–476. IEEE (2010)

23. Jain, R., Kasturi, R., Schunck, B.G.: Machine Vision. Computer Science Series. McGraw-Hill, New York (1995)

24. Jeon, S., Choi, S.: Haptic augmented reality: taxonomy and an example of stiffness modulation. Presence Teleoper. Virtual Environ. **18**(5), 387–408 (2009)

25. Kitagawa, M., Yamamoto, T.: 3D puzzle guidance in augmented reality environment using a 3D desk surface projection. In: 2011 IEEE Symposium on 3D User Interfaces (3DUI), pp. 133–134. IEEE (2011)

26. Kovac, J., Peer, P., Solina, F.: Human skin color clustering for face detection, vol. 2. IEEE (2003)

27. Krig, S.: Computer Vision Metrics: Survey, Taxonomy, and Analysis. Springer, Berlin (2014)

28. Lamberti, F., Manuri, F., Sanna, A., Paravati, G., Pezzolla, P., Montuschi, P.: Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance. IEEE Trans. Emerg. Top. Comput. **2**(4), 411–421 (2014)

29. Lehtinen, V., Oulasvirta, A., Salovaara, A., Nurmi, P.: Dynamic tactile guidance for visual search tasks. In: Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, pp. 445–452. ACM (2012)

30. Lepelley, M.C., Lejeune, L., Lestienne, F.: Study of tactile guidance of the hand in a pointing task by means of a vibrotactile device. Comput. Methods Biomech. Biomed. Eng. **12**(S1), 169–170 (2009)

31. Luces, J.V.S., Okabe, K., Murao, Y., Hirata, Y.: A phantom-sensation based paradigm for continuous vibrotactile wrist guidance in two-dimensional space. IEEE Robot. Autom. Lett. **3**(1), 163–170 (2018)

32. Mante, N., Weiland, J.D.: Visually impaired users can locate and grasp objects under the guidance of computer vision and non-visual feedback. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1–4. IEEE (2018)

33. Nee, A.Y.C., Ong, S.K., Chryssolouris, G., Mourtzis, D.: Augmented reality applications in design and manufacturing. CIRP Ann. Manuf. Technol. **61**(2), 657–679 (2012)

34. Nishihara, A., Okamoto Jr., J.: Object recognition in assembly assisted by augmented reality system. In: SAI Intelligent Systems Conference (IntelliSys), 2015, pp. 400–407. IEEE (2015)

35. Okamoto Jr., J., Nishihara, A.: Assembly Assisted by Augmented Reality (A3R), pp. 281–300. Springer, Berlin (2016)

36. Ong, S.K., Yuan, M.L., Nee, A.Y.C.: Augmented reality applications in manufacturing: a survey. Int. J. Prod. Res. **46**(10), 2707–2742 (2008)

37. Oron-Gilad, T., Downs, J.L., Gilson, R.D., Hancock, P.A.: Vibrotactile guidance cues for target acquisition. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **37**(5), 993–1004 (2007)

38. Palmarini, R., Erkoyuncu, J.A., Roy, R., Torabmostaedi, H.: A systematic review of augmented reality applications in maintenance. Robot. Comput. Integr. Manuf. **49**, 215–228 (2018)

39. Papagiannakis, G., Singh, G., Magnenat-Thalmann, N.: A survey of mobile and wireless technologies for augmented reality systems. Comput. Animat. Virtual Worlds **19**(1), 3–22 (2008)

40. Pathomaree, N., Charoenseang, S.: Augmented reality for skill transfer in assembly task. In: IEEE International Workshop on Robot and Human Interactive Communication, 2005. ROMAN 2005, pp. 500–504. IEEE (2005)

41. Salonen, T., Sääski, J.: Dynamic and visual assembly instruction for configurable products using augmented reality techniques. In: Yan, X.T., Jiang, C., Eynard, B. (eds.) Advanced Design and Manufacture to Gain a Competitive Edge, pp. 23–32. Springer, London (2008)

42. Salonen, T., Sääski, J., Hakkarainen, M., Kannetis, T., Perakakis, M., Siltanen, S., Potamianos, A., Korkalo, O., Woodward, C.: Demonstration of assembly work using augmented reality. In: Proceedings of the 6th ACM International Conference on Image and Video Retrieval, pp. 120–123. ACM (2007)

43. Sergi, F., Accoto, D., Campolo, D., Guglielmelli, E.: Forearm orientation guidance with a vibrotactile feedback bracelet: on the directionality of tactile motor communication. In: 2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics, pp. 433–438. IEEE (2008)

44. Siltanen, S., Hakkarainen, M., Korkalo, O., Salonen, T., Sääski, J., Woodward, C., Kannetis, T., Perakakis, M., Potamianos, A.: Multimodal user interface for augmented assembly. In: IEEE 9th Workshop on Multimedia Signal Processing, 2007. MMSP 2007, pp. 78–81. IEEE (2007)

45. Straub, M., Riener, A., Ferscha, A.: Route guidance with a vibrotactile waist belt. In: 4th European Conference on Smart Sensing and Context (2009)

46. Syberfeldt, A., Danielsson, O., Holm, M., Wang, L.: Visual assembling guidance using augmented reality. Procedia Manuf. **1**, 98–109 (2015)

47. Tan, H., Gray, R., Young, J.J., Taylor, R.: A haptic back display for attentional and directional cueing. J. Haptics Res. **3**, 1–20 (2003)

48. Tang, A., Owen, C., Biocca, F., Mou, W.: Comparative effectiveness of augmented reality in object assembly. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80. ACM (2003)

49. Teague, M.R.: Image analysis via the general theory of moments. JOSA **70**(8), 920–930 (1980)

50. Treiber, M.A.: An Introduction to Object Recognition: Selected Algorithms for a Wide Variety of Applications. Advances in Computer Vision and Pattern Recognition. Springer, London (2010)

51. Tumler, J., Doil, F., Mecke, R., Paul, G., Schenk, M., Pfister, E.A., Huckauf, A., Bockelmann, I., Roggentin, A.: Mobile augmented reality in industrial applications: approaches for solution of user-related issues. In: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, pp. 87–90. IEEE Computer Society (2008)

52. Tzorakoleftherakis, E., Murphey, T.D., Scheidt, R.A.: Augmenting sensorimotor control using "goal-aware" vibrotactile stimulation during reaching and manipulation behaviors. Exp. Brain Res. **234**, 1–12 (2016)

53. Valentini, P.P.: Natural interface for interactive virtual assembly in augmented reality using leap motion controller. Int. J. Interact. Des. Manuf. (IJIDeM) **12**, 1–9 (2018)

54. Van Krevelen, D.W.F., Poelman, R.: A survey of augmented reality technologies, applications and limitations. Int. J. Virtual Real. **9**(2), 1 (2010)

55. Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M., von der Heydt, R.: A century of gestalt psychol-

ogy in visual perception: I. Perceptual grouping and figure–ground organization. Psychol. Bull. **138**(6), 1172 (2012)

56. Wang, X., Ong, S.K., Nee, A.Y.C.: A comprehensive survey of augmented reality assembly research. Adv. Manuf. **4**(1), 1–22 (2016)

57. Webel, S., Bockholt, U., Engelke, T., Gavish, N., Olbrich, M., Preusche, C.: An augmented reality training platform for assembly and maintenance skills. Robot. Auton. Syst. **61**(4), 398–403 (2013)

58. Weber, B., Schätzle, S., Hulin, T., Preusche, C., Deml, B.: Evaluation of a vibrotactile feedback device for spatial guidance. In: 2011 IEEE on World Haptics Conference (WHC), pp. 349–354. IEEE (2011)

59. Wertheimer, M.: Laws of organization in perceptual forms. In: Ellis, W.D. (ed.) A source book of Gestalt Psychology. Harcourt Brace, New York (1923)

60. Wiedenmaier, S., Oehme, O., Schmidt, L., Luczak, H.: Augmented reality (AR) for assembly processes design and experimental evaluation. Int. J. Hum. Comput. Interact. **16**(3), 497–514 (2003)

61. Yuan, M.L., Ong, S.K., Nee, A.Y.C.: Augmented reality for assembly guidance using a virtual interactive tool. Int. J. Prod. Res. **46**(7), 1745–1767 (2008)