A new approach to assist Virtual Image accessibility for Visually Impaired People ⋆

Edwige Pissaloux, Lilia Djoussouf, Katerine Romeo, Ramiro Velazquez, Simon Gay, Ngoc-Tan Truong, Son Dao

A new approach to assist Virtual Image accessibility for Visually Impaired People*

Edwige Pissaloux[1[0000000198912531]], Lilia Djoussouf[1[0000−0001−5204−832X]], Katerine Romeo[1[0000000161756248]], Velazquez Ramiro[2[0000−0001−9966−9131]], Simon L.Gay[3[0000000271071055]], Ngoc-Tan Truong[1[0000000177375773]], and Son Duy Dao[1[0000−0002−5253−6398]]

[1] LITIS, University of Rouen Normandy, Rouen, France
edwige.pissaloux@univ-rouen.fr
[2] Universidad Panamericana, Aguascalientes, Mexico
[3] LCIS, University of Grenoble Alpes, Valence, France

Abstract. Currently, graphical data is becoming increasingly ubiquitous with new technologies. However, today's technologies are still of limited access to such representations (images, graphs, charts. . .) for Visually Impaired People (VIP). The quantity and quality of presented information are key points of efficient accessibility. Therefore, this paper proposes the presentation of such information via a tactile gist (a tactile representation of essential data). New rules for 2D data representation (e.g. paintings, images, maps) are proposed via the tactile gist, which helps us, especially VIP, to understand them. These rules are deduced from experiments lead with VIP for tactile gist representations on two supports – thermoformed paper and dedicated original force-feedback based device named F2T (Fore Feedback Tablet). These rules take into account the human touch/haptic sense specificities and human cognitive capabilities. Such rules should be included in the design of any assistance to 2D data accessible by the VIP.

Keywords: tactile gist representations· 2D data accessibility · Force-Feedback Tablet (F2T)· virtual images· perception· Visually Impaired People (VIP)

1 Introduction

Nowadays, culture, education, and many daily life tasks are getting increasingly reliant on 2D data such as pictures, schematics or maps. However, accessing such information is still difficult for Visually Impaired People (VIP). Traditional

* This research has been financially supported, via different research projects, by the CCAH (Le Comité National Coordination Action Handicap), Région de Normandie, and European Commission (FEDER) and ANR (Inclusive museum Guide). We thank students and staff of the University of Rouen Normandie, and the VIP, members of our research team or from VIP charties, for their participation in our evaluations.

screen readers and braille devices are not adapted to convey 2D data efficiently, limiting VIP's inclusion in society and lowering the quality of their life.

The access to the culture, especially to the artworks presented in museums, designed mainly for sighted people, are not accessible to VIP. In traditional approach (provided e.g. by classic museum guides), artworks are oculo-centred representations accessible for the sighted public. The VIP accessibility is considered thanks to a transcription of some visual information into audio descriptions; sometimes, static tactile representations [1] are also available. Figure 1 gives two examples of artworks' presentations, which target the VIP, from the British Museum (Fig. 1a and 1b), and the Fine Art Museum, Rouen, France (Fig. 1c and 1d). Both examples use the same technology (namely, thermo-inflation) to fix a 3D scene on a 2D support. Figure 1a presents a scene from the Greek temple in relief with its fronto-parallel and orthographic projections (Fig. 1b). Figure 1c presents Sisley's Seine River painting and its fronto-parallel projection (Fig. 1d). Both representations try to make a "tactile copy" of the observed 3D scenes, which are not relevant for tactile/haptic explorations. Indeed, these representations are overloaded with details, which imply long exploration time and high cognitive load. Moreover, the details' superposition does not allow human fingertips [2] to extract the meaning from the object's contours (edges). Furthermore, the physical nature of the supports does not allow to implement the scaling operations (zooming in and out), as those presentation are unchangeable (frozen). Therefore, such oculo-centric-tactile representations cannot be easily perceived by VIP.

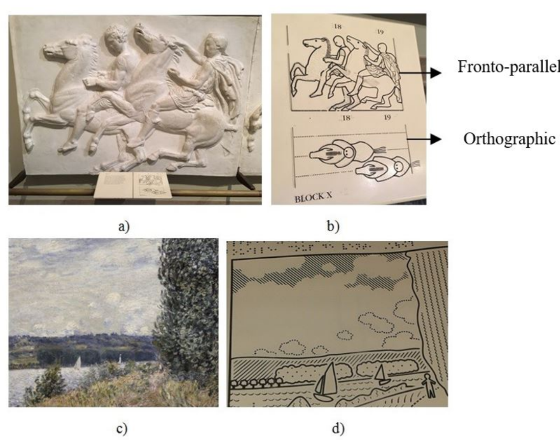


Fig. 1. Static tactile representations a) The moulded friezes of a Greek temple, British Museum ; b) fronto-parallel and orthographic projections of relief from a); c) Boats on Seine by Sisley and d) their thermoformed representation.

However, the artworks can be perceived in a de-visualized way[1], i.e. perceived without vision (or only with vision as a secondary sense) in a multimodal way. Such a presentation will create new aesthetic and emotional impressions, deepen our knowledge of the creative gesture and lead to the effective inclusion of all audiences to digital society. Therefore, considering the de-visualization principle a new generation of museum guides will offer new experiences of artworks. We propose its integration via a simulation of natural cognitive processing (i.e. via an artificial intelligence cognitive process) which underpins concepts/percepts emergence from stimulations following sensory-motor theory of visual perception [3].

Therefore, this paper proposes a new approach to design representations of 2D data accessible for VIP, based on the concept of "a tactile gist presentation" (Section 2). A set of rules, deduced from experiments lead with VIP, is proposed, allowing the creation of haptic (physical and virtual) representations of images on different supports, such as thermoformed paper or an original force-feedback based haptic device, named F2T (Section 3). Some examples of created virtual images, which can be explored by any public via tactile or haptic assistance, are provided in Section 4. Section 5 encompasses some future works.

2 Gist Concept

Visual gist is a global minimal information presenting an image where usually only a few objects of important size are perceived [4, 5]. The gist representation can be seen as a global and rough "map" of an image as it allows to grasp a minimal meaning of the image in a single glance [6]. Therefore, the role of gist presentation is threefold: (1) representing an image in a synthetic way, (2) allowing the relative localization of objects in an image, and (3) assisting the decision for further behavior (e.g. analyzing the image in more details or changing the image). In function of the support, we speak about the visual or tactile (haptic) gist.

An example of the (visual) gist representation is shown in Fig.2; its left image (a) is the scene of "Raven and fox" extracted from the Bayeux Tapestry, XI century (France, [7]), while right image (b) represents its associated (tactile) gist where 4 circles delimit four important areas each having a specific cognitive meaning.

3 Rules for Creation of Haptic Representation of Images Accessible to VIP

The rules for creation of haptic gist representation have been established thanks to a series of experiments with end-users [8–11]. To create VIP experience of haptic images, a specific evaluation procedure has been defined (cf. Fig. 3). In our case, the gist representation is implemented based on two supports: thermoformed paper (A path) and the original force-feedback tablet (B path). The (A)

Fig. 2. a) Example of Bayeux Tapestry “Raven and fox”, its gist representation of a 2D image [7].

.

path allows the perception, thus the evaluation, of the static elements of tactile gist (e.g. a tree, a cheese). The (B) path, thanks to force-feedback tablet F2T providing haptic and kinesthetic representation, allows the perception (feeling), and the evaluation of static and dynamic elements of an image (e.g. to follow object edges or the speed of the dropped cheese).

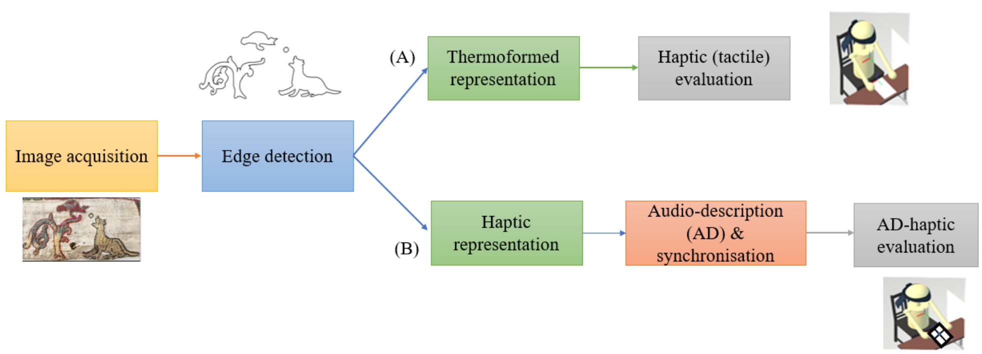


Fig. 3. Procedure for haptic image accessibility evaluation.

Early evaluations, with VIP or blindfolded participants, of different images’ gist representations [8–11] allowed establishing the following rules and definitions of tactile/haptic representations of images:

Rule 0. A haptic gist presentation of an image (painting) should be built as this image surjective projection.

Rule 1. The tactile gist consists of a set of regions having an important meaning for considered images.

Rule 2. A gist of an object is defined by its bounding box (simplified and approximated).

Rule 3. An ad hoc, but cognitively useful, simplification and/or approximation of the haptic gist of an image are mandatory.

Rule 4. A bounding box of a gist should be clearly spatially isolated from others to extract its semantic meaning during the manual exploration (spatial separations between boxes should be introduced, if necessary).

Rule 5. The tactile-only representation is relevant for isolated simple objects' representations, close to known geometric figures.

Rule 6. The tactile-only representation is insufficient for complex figure comprehension.

Rule 7. The tactile gist should be completed by an ad hoc audio-description.

Thermoformed and haptic representations of images by tactile gist are shown in more detail in Section 3.1 and Section 3.2, respectively.

3.1 Thermoformed Representation of tactile gist

Gist representation using thermoformed papers can be used to help VIP to perceive a static scene or image with their fingers. By applying the above rules, the final tactile gist for Fig. 2a, which has been validated with VIP and blindfolded students, is shown in Fig. 4. It should be observed that the initial relative positions between objects of the scene have been enlarged by adding the additional space for fingertips exploration (cf. Rule 4). This tactile gist representation can be directly printed using a thermoforming printer.

The thermoformed representations allow us to create representations of the scene containing the static elements only; they may help to find the most relevant approximations and simplifications. However, the tuning of their representations is time-consuming and complex, and their effective realization requires a dedicated printer. Being mono-modal, their comprehension may be difficult when manually explored. Moreover, it is impossible to change the thermoformed display because each image requires its thermoformed representation.

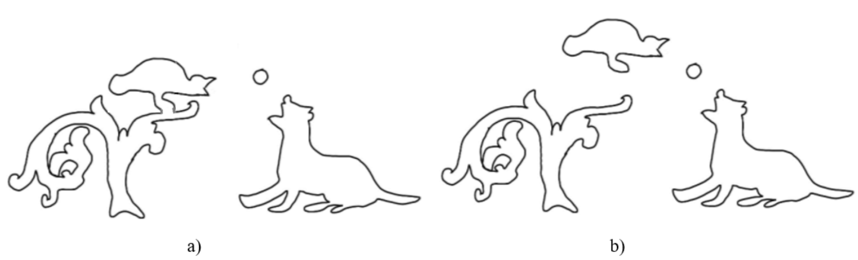


Fig. 4. Initial (a) and final (b) tactile gist presentation of the scene "Raven and fox" of Bayeux Tapestry.

A force-feedback tablet can overcome such limits and it allows accompanying the tactile/haptic representations by an audio description lead to a TAD (Tactile-Audio-Described) representation.

3.2 Haptic effects for gist representation

The Force Feedback Tablet (F2T, Fig. 5a) is a new architecture, based on the force-feedback principle, to generate haptic effects associated with visual cues used by painters such as height/depth variations, borders (edges), and textures. The F2T acquires the user's movement intentions via the thumbstick, independently from the current end-effector movement (Fig. 5b). This specificity allows the implementation of a wide variety of dynamic and interactive haptic effects useful for 2D data representations. Some of them have been already provided by the F2T (edge and relief, flow, rail, attractors, and active guidance); others are still under investigation.

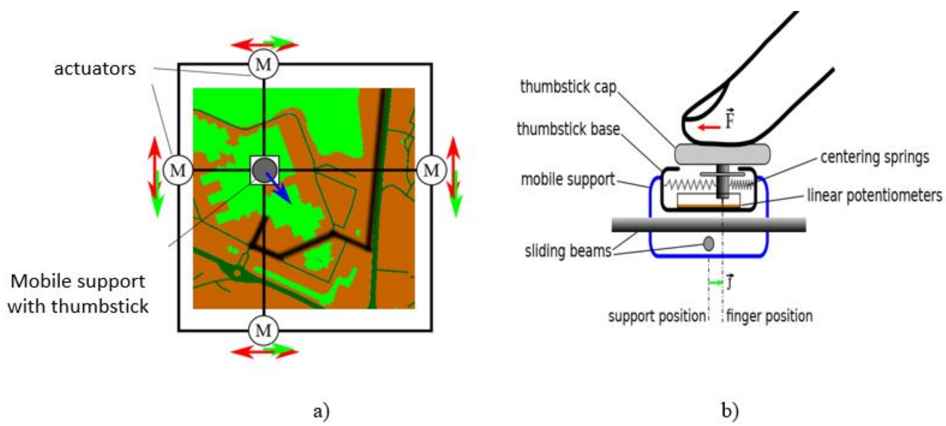


Fig. 5. a) F2T architecture (attached to a screen on which a map is displayed); b) thumbstick architecture which local movement depends on the image content at (x,y) [11].

F2T principle is presented in Fig. 6. The initial intended speed $V^{'}$ is the movement generated by the device to follow the user's finger. The output speed $V^{'}$ is modified by haptic effects that are applied to the mobile support, and P(x,y) is the current position of the mobile support on the available surface (screen attached to F2T). The force-feedback principle allows producing several haptic effects, which add offsets to the control speed $V^{'}$. These effects can be cumulated by applying them successively. In this case, the output speed $V^{'}$ of an effect is used as the input speed vector V of the next effect. The number and order of applied effects can be changed according to the desired haptic perceptions to induce.

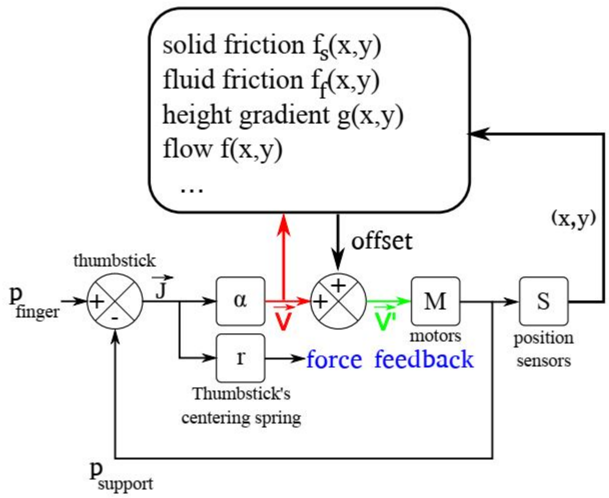


Fig. 6. Principle of different effects generation at (x,y) (frictions, flow, etc.) [11].

In our first academic prototype (Fig.7), the following effects were implemented: edges, and reliefs, attractors, flows, railings, and waves (sine waves simulated). The guidance mode of the device is also presented.

Edge and relief effects allow representing objects' edges. When the edge gradient is small (the slope of the object's edge is small), the user's finger will be either slowed down or accelerated depending on the direction of the edge gradient. This effect can be used to explore a bas-relief or to guide the users along a trench, maintaining their finger on a path. When the edge gradient is strong enough, the movement will stop. In that case, the user's finger can slide along the object's edge which speeds up the object's shape recognition.

Flow effect adds an offset to the finger's movement. The movement offset value and direction can be defined at each point of the surface. It is thus possible to represent dynamic elements with mobile parts (e.g. vehicles or conveyors) or dynamic physic elements (e.g. wind in Fig. 8a).

The rail is an effect with no equivalent in the real world (e.g. Fig. 8b). This effect is based on a predefined a rail curve: movements following the rails will not be affected, while movements crossing the rails will be slowed down (reduction of the component of the movement that is perpendicular to the rails). This effect adds constraints to the user's movements.

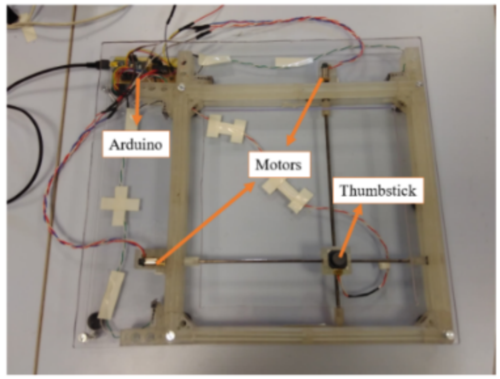
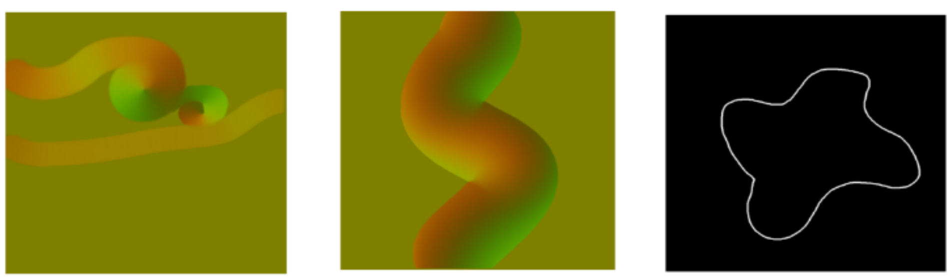


Fig. 7. F2T prototype [11].

Attractors are areas or points generating an attractive or repulsive force on the user's finger (Fig. 8c). The force strength, profile, and range can be predefined for each attractor. This effect can be used to simulate magnetic paths, holes, bumps, or even gravity fields.



a) b) c)

Fig. 8. F2T original haptic effects: a) and b) images encode vector fields that can be used for flow and rail effects, c) attractor image.

After presenting the F2T architecture, Section 4 introduces a case study of a real scene using F2T and outlines the principles of the guided mode applying the tactile gist concept.

4 Case Study

For the gist, we can also say that it is minimal information to represent the content of the message of an image. Based on this concept, F2T presents two modes: free exploration and guidance. In free exploration mode, the user can move the thumbstick anywhere on the image. Depending on the thumbstick position, the sound of a corresponding dynamic element will be activated. That helps VIP understanding the general content of the image. Adding more experiences accessing the painting for VIP, F2T proposes the second mode – guidance mode. This mode is powerful; the user does not move the thumbstick. This mode is programmed by a sequence of way points defining the path. This path is built to present the important elements. The thumbstick will be moved automatically following this path. Moreover, the audio-descriptions are also programmed for the corresponding elements (multimodal presentation). Section 4.1 and Section 4.2 present the details of the two modes

4.1 F2T free exploration mode

When creating a gist representation of a considered image displayed by the F2T, the haptic effects introduced above are applied to different parts of the image. Figure 9a represents scene 24 of Bayeux Tapestry [7]; Figure 9b delimitates four-regions gist where the red region corresponds to the boat and the sea, the yellow region matches the shields placed on the edges of the boat; the dark green region indicates the sail, and the light green region matches the wind (air) which pushes the sail. When the finger is in a specific region, the sounds of the corresponding elements can be activated. The rail effect is applied to the dark green region and will guide accordingly the finger (on the thumbstick) through its shape and with the associated speed (Fig. 9c). The wave of the sea are simulated with a sinusoid function; a sea wave sound is activated when the thumbstick is in the low red region (figure 9d). VIP can freely explore this gist representation on F2T.

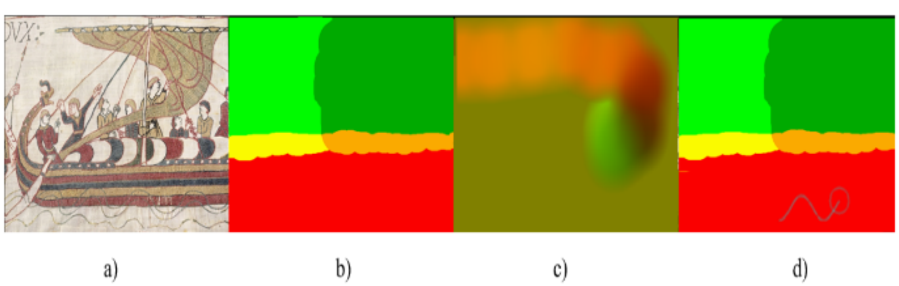


Fig. 9. a) Scene of Bayeux Tapestry [7], b) a gist representation of the image in a); rail effect c) simulates the wind; wave effect d) simulates (via a sinusoid) the wave of the sea.

4.2 F2T guided mode

Besides the free exploration mode (corresponding to VIP natural exploration of tactile representations), F2T provides the guided mode. The guidance assists a detailed exploration of an element of a gist by following its outline. The object's outline is thus represented by a curve (a path) defined by a list of line segments (salient points of the object). Figure 10 gives an example of a path while following the shape of the raven (the pink points). The points of segments are memorized in a text file with attributes of the guidance such as the point coordinates, the exploration speed at this point, the attraction force, etc. Using this file, the F2T control software (Fig.11) generates a virtual representation of the image that the user can explore in a controlled way.

Fig. 10. The suggested exploration path (pink points) following the raven's shape.

A simple script language has been developed to load or change images and paths, and to play sound files when predefined conditions are fulfilled (cf. Fig.11 right). This scripted system makes it possible to define scenarios and synchronize the guidance with audio-descriptions (multimodal presentation), or to change the virtual environment configuration (e.g. change the image or user's profile). Figure 11 presents the guidance mode for "Raven and fox". The paths defined by green points correspond to the four important elements of the scene (tree, raven, cheese, and fox), and constitute a scene overview. An audio-description may accompany the haptic exploration of each element.

Such multimodal guide mode will help VIP to recognize the contours of elements and understand the whole scene.

5 Conclusion

This paper proposes an ICT (Information and Communication Technology) system to assist the de-visualization of artworks. Such accessibility is based on the concept of multimodal presentation of artworks with the gist concept. The gist representation allows building "a map of a scene or image" (the gist of level 1)

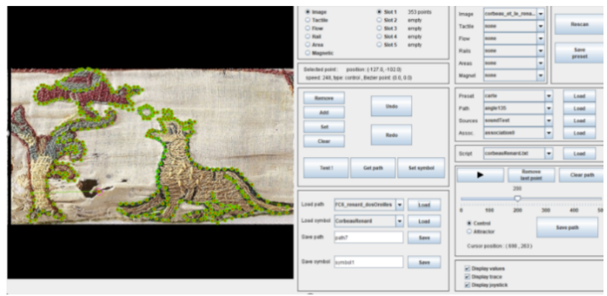


Fig. 11. Guidance mode control panel on F2T interface (green points corresponding to the whole picture guided exploration).

and can provide a more detailed representation of each element of the “map” (the gist of level 2). Besides the gist representations, new rules for gist implantation on different hardware supports (namely, thermoformed paper and force-feedback tablet) have been established based on experimental evaluation of haptic representations with VIP and blindfolded participants. The thermoformed presentation, although creating faithful representations of the painting, may assist the perception of simple static elements only. The multimodal force-feedback device F2T presented here briefly overcomes this limit as it provides the multimodal tactile-audio-described representations. F2T can generate some haptic effects such as height/depth variations, borders, and textures. This device helps VIP to access 2D data (maps, graph, 2D art. . .); this paper presented two modes: free exploration and guided mode to access the paintings’ different elements. In near future, a new F2T version will be designed as a frame to be clipped on classic screens; the image displayed on a screen will be used by the F2T control system. We will also investigate more haptic and audio effects to generate interactive multimodal representations of images which will be evaluated with VIP.

References

1. Thompson, H., Blindness Arts (co-ed. V. Warne), Disability Studies Quarterly, vol. 38, no.3, 2018
2. Johnson., K.O., The roles and functions of cutaneous mechanoreceptors, Current Opinion in Neurobiology, 11(4): 455–461, 2001.
3. O’Regan J.K., Noë A., What it is like to see: A sensorimotor theory of visual experience , Synthèse, 2001, 129, 79-103
4. Potter, M. C. (1976). Short-term conceptual memory for pictures, J. of Experimental Psychology: Human Learning and Memory 2, 509–522.
5. Oliva, A., Gist of the Scene, Neurobiology of Attention, Elsevier, 2005, pp.251- 258.
6. Pissaloux E., Velazquez R., Maingreaud F., A New Framework for Cognitive Mobility of Visually Impaired Users and Associated Tactile, Device, IEEE Trans. Human-Machine Systems, vol. 47 Issue 6, pp.2168-2291, Dec 2017

7. https://www.bayeuxmuseum.com/en/the-bayeux-tapestry/ [Last Accessed July 2021]
8. Romeo K., Chottin, M., Ancet, P., Pissaloux, E., Access to Artworks and its Mediation by and for Visually Impaired Persons, ICCHP 2018, July 9-13, 2018, Linz, Austria, LNCS II 233-236
9. Romeo K., Chottin, M., Ancet, P., Pissaloux, E., Simplification of Painting Images for Tactile Perception for Visually Impaired Persons, ICCHP 2018, July 9-13, 2018, Linz, Austria, LNCS II 250-257
10. Velázquez, R., Pissaloux, E., Del-Valle-Soto, C., Arai, M., Valdivia, L.J., Del Puerto, J.A., Gutiérrez, C.A., Performance Evaluation of Active and Passive Haptic Feedback in Shape Perception, 2019
11. Gay, S., Pissaloux, E., Romeo, K., Truong, N.T., F2T: A Novel Force-Feedback Haptic Architecture Delivering 2D Data to Visually Impaired People, IEEE Access, page(s): 1-11, Print ISSN: 2169-3536, Online ISSN: 2169-3536, https://doi.org/10.1109/ACCESS.2021. 3091441.