

An enactive based realtime 3D self-organization system for the exploration of a cultural heritage data base

Landy Rajaonarivo, Eric Maisel, Pierre De Loor

Lab-STICC

ENIB/UBL

Centre Européen de Réalité Virtuelle, Brest, FRANCE

{rajaonarivo, maisel, deloor}@enib.fr

Abstract—This paper presents a first step in the realization of an interactive user interface that organizes itself according to the user exploration of a database of cultural heritage objects. The first part makes a brief related works and lays the basis of this kind of system according to the enactive paradigm. The second part explains the mechanisms underlying the self-organization of the interface: keywords and cultural heritage objects are 3D graphical entities endowed with autonomous behaviors. They share a common virtual environment. Keywords behaviors are based on boids flocking simulation while cultural heritage objects appears in a virtual museum which evolves and grows progressively. The result that is presented, is an incremental construction of an interactive and realtime 3D metaphor of virtual museum which is then user-specific.

Keywords-information visualization, data exploration, self-organization, enactivism, visual metaphor

I. INTRODUCTION

The objective of this paper is to present first results on an original interface that allows the exploration of a database of cultural heritage objects. The goal is to improve the understanding and the exploration of this database content, thanks to an on-line self-organization of the graphical representation, according to the user's activity. This study mixes adaptive methods of information visualization and self-organization of data while introducing the user in the interactive loop. Different studies try to adapt presentations of information to a user [5]. For instance, [1][17] are concerned with the auto-organization of 3D data by the use of Kohonen maps or flocking algorithms [8][11][13] inspired by the seminal work of Reynolds [12]. These algorithms organize the data according to some similarities between them and show shapes or graphical structures which help the user to understand the data. However, these algorithms do not take into account the intents of users or their strategies of exploration. At the opposite, approaches like [2],[14] propose an adaptation of representations to a user profile, which is elaborated from his traces of interactions. The first one is based on the notion of degree of interest mixed with an ontology of the data, while the second one elaborates a dynamic Bayesian networks. One of our objective is to gather all this ideas to propose a virtual environment called *enactive* as it is inspired by an enactivist argumentation.

Enaction is a field of cognitive science, rooted in biology and close to constructivism [6], [16]. It proposes an explanation to the phenomena of sense-making which is not addressed by cognitivism [15]. According to enaction, the basis of the understanding, relies on the history of the interactions between an individual and his environment. For the present work, two points inspired by enaction are raised: the first one is the importance of the link between actions and perceptions for cognition. Humans find a stable representation of their world from sensorimotor invariant [7], [10]. Following these principles, some studies tried to develop the notion of enactive systems. For instance, [4] propose an enactive media in which the function of interacting is driven by bodily involvement and spatial presence of the human agent without the assumption of conscious control of the system. The second point, which is more important for this present study, is the coupling between the environment and the individuals: humans modify their environment which, in return, impacts on the possibility of actions for humans. After a while, this co-construction of the agent and his environment implies a strong interdependence between them. This coupling is inseparable to the explanation of the cognition and of the representations of individuals.

One important point is that each couple agent/environment leads to a specific coupling that let emerge a specific understanding, individual to each person. The challenge here, is to propose a virtual environment which favor the emergence of understanding. We aim at defining an intelligent system designed to help a human to construct a representation of a problem with it. Then, the previous coupling principles transform the idea of artificial intelligence based of predefined knowledge (classical artificial intelligence) toward an evolving, not predefined and human coupled system that is an enaction-based artificial intelligence system [3].

In our studies we propose a 3D realtime self-organized environment, implemented as a virtual museum, that allows the co-construction of the environment and of the user meaning. This evolution depends on the users interactions and on links between cultural heritage objects and keywords. Consequently, the museum become the user's museum which retrace his exploration as a reflexive trace of his

understanding.

The next part (Section II) of this article presents the main principles of our proposition and highlight the two main features which constitute its originality respectively in Section II-A and II-B. Section III illustrates the result of the implementation and then, Section IV presents some perspectives of this study.

II. PRINCIPLES

General principles of our database exploration system is depicted by figure 1. Two types of representations are presented to the user: 1) some keywords extracted from the database according to a textual description of each cultural heritage objects are positioned in a *keywords space*, 2) the corresponding images of cultural heritage objects arranged on shelves in rooms of the virtual museum.

The spatial organization and the temporal evolution of the objects and of the rooms of the museum are ensured according to algorithms which take into account user's interactions. These different animations highlight ideas or concepts dynamically which can influence the user. In fact, through interactions, he indicates his preferences either on keywords or objects. Clicking on objects, moving an *marker of interest* in the *keyword space* (see below) or keeping ahead of objects are indications of the user's interest. Beyond this graphical interface, there are statistical links between keywords and also between keywords and cultural heritage objects. A recommendation system computes these links from the textual description of objects in the database. This recommendation system is not described here. However, we consider it as a black box which allows the construction of a recommendation graph. This latter is used in order to present the statistic links between the nodes (keywords or objects) which are returned by the recommendation system.

Let $G_r = (N_r, \Gamma_r)$ the graph such as:

$N_r = Ks \cup Os$ a set of nodes with Ks is a set of keywords while Os is a set of objects

$\Gamma_r = \{(n_1, n_2) \mid n_2 \text{ is recommended by } n_1\}$ a set of the edges which represents recommendations.

$V^-(n) = \{k_i \mid (k_i, n) \in \Gamma_r, k_i \in Ks, n \in N_r\}$ is a set of keywords which recommended the node n

The graph G_r is completed gradually for each call of recommendation system.

Then, the temporal evolution is computed by steps as following: for each step, the *most interesting keywords* for the user can be defined from his interactions. These keywords are introduced to the recommendation system for the purpose of adding new keywords in the *keyword space* and new objects in the virtual museum. Keywords move in front of the user (see Figure 2): most supposed interesting keywords are approaching the *marker of interest* when less interesting move away, toward a *marker of disinterest*. In parallel, new objects (provided by the recommendation system according to the most interesting keywords) appear on the shelves and

new rooms are constructed close to the user. All actions of the user on elements of the environment called "user actions history" allow the system estimate the user representation: his preferences and his disinclination on some keywords or some heritage objects. Then, the user representation is used to reinforce the pertinence of the keywords. So, its behaviors in its environment. As a result, the museum grows progressively, but as it grows, it proposes new possibilities of interactions, new affordances for the user. It favors the enactive principle of a co-construction. After a while, the museum become specific to the user and to the history of his interactions.

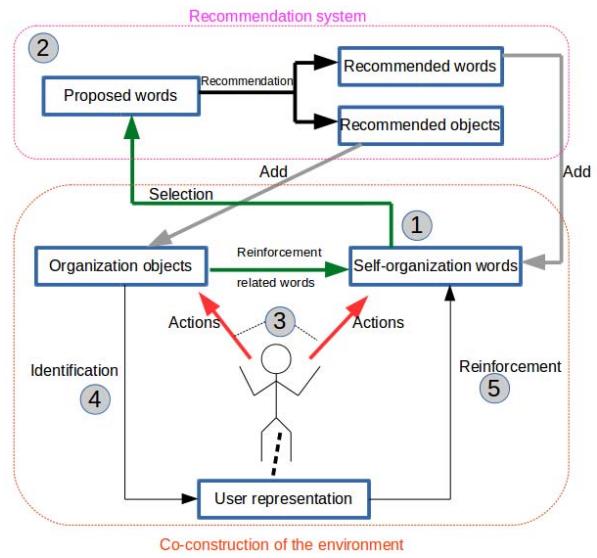


Figure 1. Principles of the system

A. Self-organization of keywords

The *space of keywords* is represented Figure 2. Keywords are represented by 3D shapes with different colors indicating their degree of interest (green: more interesting keywords, yellow: averagely interesting keywords and red: less interesting keywords). Each keyword adopts a behavior which depends on the group of interest for which it belongs. As keywords in a group have the same behaviors, they gather. This formed groups could be meaningful to the user. Each period of time, the pertinence of all keywords are re-evaluated, as well as their behaviors. These behaviors keywords are represented by the forces applied to it. The blue arrows in figure 2 show the different forces applied on one keyword and the red arrow represent the resultant force. Their displacements are inspired on booids and information flocking [11], [9] for which spatial elements interact together thanks to forces of attraction or repulsion and to birth-death mechanisms. The force that control the displacement of spatial keywords is computed from three elements: 1) the behavior of the user, 2) the two markers

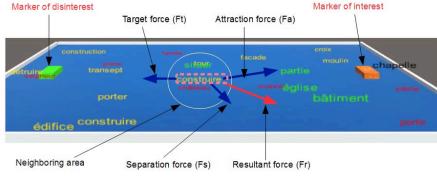


Figure 2. Representation of the keywords space: the colors green, yellow, and red are respectively the colors of the keywords in S_{ks} , M_{ks} and W_{ks}

(of *interest and of disinterest*) and 3) the recommendation link of the graph G_r . In order to describe the computation of these forces, let K_s be the sorted set of keywords in the 3D *keywords space*. At any time, one of these keyword, named w^* is considered to be the more interesting for the user (we will explain latter how this keyword is designed). K_s is structured into 3 subsets $K_s = S_{ks} \cup M_{ks} \cup W_{ks}$, when S_{ks} , M_{ks} and W_{ks} correspond respectively to the subsets of keywords with strong interest (including w^*), with moderate interest and with weak interest. Most interesting keywords (S_{ks}) are attracted by the target *marker of interest* represented by an orange cube on figure 2. This cube can be moved by the user to indicate some words which have his preference (whatever the content of S_{ks}). The less interesting keywords (W_{ks}) are attracted by the target *marker of disinterest* represented by a green cube on figure 2. The force of attraction of these two targets is represented by \vec{F}_t . Each word w in K_s has a set $V^-(w)$ which contains the list of other keywords that cause their apparition in K_s after they was sent to the recommendation system. Only the words in M_{ks} are attracted by their parents with a force \vec{F}_a . Each word w in K_s has also an existence time $A(w)$ (in seconds) since its apparition in K_s and an interest value $I(w)$. $A(w)$ represents the realtime age of w when $I(w)$ represents the interest of the user, which also evolves in realtime. A separating force \vec{F}_s allows to avoid collision between keywords during their movements in order to preserve the readability of the graphical representation. Eventually, the force which defines the movement of 3D keywords is $\vec{F} = \alpha * \vec{F}_t + \beta * \vec{F}_a + \gamma * \vec{F}_s$ with $\alpha + \beta + \gamma = 1$

Three process run in parallel :

- 1) A simulation loop manages the movements of the keywords as well as their reclassification (S_{ks} , M_{ks} , W_{ks}) in realtime:

- We limit the number of keywords in K_s , in order to preserve the readability of the graphical representation. Let N_{max} the maximal number of keywords in the "keywords space". At each time step, all keywords in (K_s) are sorted according to their pertinence $P(w)$. If the number of the ks elements is more than N_{max} , then the N_{max} most relevant keywords are conserved and the rest (less relevant) are removed from the *keywords space*.

- Let N be the size of the sorted set K_s . This set is divided into the 3 subsets S_{ks} , M_{ks} and W_{ks} which contain respectively the first $N/3$, the second $N/3$ and the last $N/3$ words of K_s .

- The different forces of displacement of the keywords are updated.

- $P(w)$ is evaluated from three elements: 1) $D(w)$, the distance between w and the marker of interest, 2) $A(w)$, the age of w in the *keyword space* and 3) $I(w)$, the interest value of w . This latter represents the interest of the user for w from his interactions. The more w is close to the marker of interest, recently added in K_s (low value of $A(w)$), and considered interesting for the user, the more it is pertinent:

$$P(w) = a * \left(\frac{1}{1+A(w)} \right) + b * \left(\frac{1}{1+D(w)} \right) + c * I(w) \text{ with } a+b+c=1$$

- A new w^* can be elected from three possible causes: 1) The increasing of the pertinence $P(w)$ due to the actions of the user, 2) The fact that a new word is close to the *marker of interest* (we define a minimal distance for that) through its motion or the displacement of the *marker of interest* by the user (which modify also $P(w)$ through the change of $D(w)$), 3) the interest of the user for an object represented in the virtual museum, which reinforces the interest of the keywords linked with it.

- 2) A callback function is triggered for two possible reasons: a) w^* changes because of the realtime process, b) user selects an object.

- a) If w^* changes, it is sent to the recommendation system. This latter gives a sorted list of recommended keywords. This sorting corresponds to a statistical occurrence of these keywords close to w^* in the texts which describe the cultural objects. The five first keywords are added at the beginning of the sorted set K_s (we limit the number of added keywords to five to favor the readability of the graphical interface, but different limits could be used). Then, for each word w which are newly added in K_s : $V^-(w) = V^-(w) \cup w^*$. It indicates that the keyword w^* is at the origin of the apparition (or re-appearance) of w in K_s .

- b) Let O_s the set of objects in the virtual museum. Each heritage object o in O_s has parents noted $V^-(o)$. This latter contains the keywords which recommended o in the museum. This technique of saving parents allows to increase brutally the interest of all keywords included in $V^-(o)$ when o is selected by the user.

$$V^-(o) = V^-(o) \cup w^*.$$

The interest values $I(w)$ of all w in K_s are updated by using a mechanism of contribution-redistribution. The idea is to keep more or less invariable the sum of all interest values and to balance the interest values between the pertinent and non pertinent keywords. Let τ the rate of contribution

i) Contribution:

$$C = \sum_{w \in K_s} (\tau * I(w))$$

ii) Redistribution:

$$R = \frac{C}{|V^-(o)|}$$

Then, $\Delta I(w) = R - \tau * I(w)$, if $w \in V^-(o)$
 $\Delta I(w) = -\tau * I(w)$, if $w \notin V^-(o)$

- c) Then, the system let the keywords move and wait that a new w^* is elected by the realtime process. This election comes from three possible causes: 1) The decreasing of the pertinence $P(w)$ due to its age, 2) The fact that a new word was nearest the *marker of interest* through its motion or the motion of the marker by the user (which modify also $P(w)$ through the change of $D(w)$, 3) the interest of the user for an object represented in the virtual museum, which reinforce the interest of the keywords linked with it.
- 3) An amortization calculation updates sequentially (synchronised on the realtime) the reassessment of the degree of interest of user on keywords in order to converge their value toward the average. We use again the technique based on the mechanism of contribution-redistribution to amortize the interest values. Let σ the rate of amortization and I_{avg} the average of interest of keywords in K_s :

a) Contribution:

$$C = \sum_{w \in K_s} \sigma * (I(w) - I_{avg}) \text{ such as } I(w) > I_{avg}$$

b) Redistribution:

$$R = \frac{C}{|\{w | I(w) < I_{avg}\}|}$$

Then, $\Delta I(w) = -C$, if $I(w) > I_{avg}$
 $\Delta I(w) = R$ if $I(w) < I_{avg}$

This keywords behavior highlights the user visually on new possibilities of researches close to his interest. Moreover, this highlight evolves in time because *old* keywords disappear to let their places to new ones. Apparitions and disappearances imply new dynamics of all the words and favor serendipity. The user can also drastically influence the result.

B. Self-organization of the virtual museum

The enactive system proposes an incremental construction technique of the museum according to the actions of the user.

For that, the construction process is called after each call of the recommendation system (step 2 of the callback function of section II-A). To do that, the virtual museum is modeled by a matrix where each cell represents a room.

The principles of co-construction of the museum is the following:

- 1) Let R the set of rooms in the museum and r one room in R having a position (i, j) in the matrix where i is its index of line and j its index of column. Each room r in R is attributed two flags : 1) r_O which defines the state of the room r . Its possible values are: *occupied* by cultural heritage objects or *empty*, 2) r_U which determines if the user is present or not in the room r .
 - 2) Let N_r the set of neighboring rooms of r which are formed by the four rooms toward which the user can access directly from r .
 $N_r = \{r(i,j-1), r(i-1,j), r(i,j+1), r(i+1,j)\}$
 - 3) In the simulation loop, we assign a qualitative value noted Q_r for each non occupied room. The possible values of Q_r are: *excellent*, *good*, *medium*. This value depends on the fact that the room is close to occupied rooms and close to the user. It's setting is based on a cellular automaton algorithm. For that, let:
- N_r^o the set of neighbors of r which are occupied.
 $N_r^o = \{N_r | r_O = true\}$
 - N_r^u the set of neighbors of r where the user is present $N_r^u = \{N_r | r_U = true\}$

So, the value of r_v is given by the following assumptions:

- if $|N_r^o| > 0$ and $|N_r^u| > 0$ then $Q_r = \text{excellent}$
- if $|N_r^o| > 0$ and $|N_r^u| = 0$ then $Q_r = \text{good}$
- if $|N_r^o| = 0$ and $|N_r^u| > 0$ then $Q_r = \text{medium}$

Each request to the recommendation system is followed by the choice of an empty room, the filling of this room with new objects and its assignation to a theme which is the input keyword w^* . The system choose in a preferential basis the empty room having an excellent value, if it does not exist, it will choose a room having a good value or, by default, a room having a medium value. After choosing the list of rooms with the better value, the system choose the room nearest to the room where the user is present.

III. IMPLEMENTATION

The virtual museum with the *keyword space* are in a unique scene showed by the figure 4. For the moment, the user interact with the elements presented in the virtual environment by using the mouse and the keyboard: clicking some heritage objects, moving keyword or marker, moving in the museum.

Heritage objects are grouped in rooms and each group is marked with a theme. The rooms are interconnected through the doors and the distance of similarity between the themes

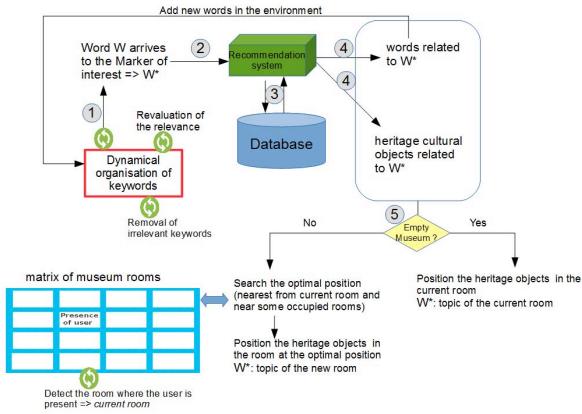


Figure 3. Co-construction of the museum



Figure 4. Exploration of database: metaphor of the museum

of two given rooms is evaluated from the minimum number of doors to cross starting from the one to get to other. So, two rooms separated only by one door have approximately similar themes. The similarity measure depends on the user intent, more specifically on the user's actions history. The theme of the four rooms connected directly with the room where the user is located, are displayed on top of the doors in order to influence the user to explore other heritage objects which are more or less similar to his intent. Gradually, as the user moves through the museum and interacts with the heritage objects or with the keywords, the museum will expand by adding progressively new rooms close of the user location. The user preferences is defined from the form of the museum thus constructed or from the trace of navigation which can be interpreted by the relationships that maintain among the themes of the rooms visited.

To establish this virtual museum, we used Unity 3D as tool and C# as language of development. We have 25 000 heritage objects in our database which are provided by a digital publishing company and are accessible on the Internet

(for anonymisation, we don't give the url in this version of submission). For our proposition, one room can contain 32 objects. So, we will need 782 rooms to present our data. In order to not get lost in the large amount of informations, our proposition allows the construction gradually according of the interest of user. The incremental filling of the rooms influence the user to visit other rooms which contain heritage objects more or less close of his interest. This filling technique evolves with times because it take into account the keywords behaviors and the user interactions.

IV. CONCLUSION AND PERSPECTIVES

This paper has described a 3D metaphor for the presentation of information in order to explore complex databases with an application to cultural heritage. Its originality lies in the use of the enactment paradigm as a conceptual basis. The enactive system is active and participates with the user, thanks to autonomous behaviors, to a co-construction of a 3D representation of the information. By this way, the result is specific to the user's exploration but is also motivated by the system. An implementation was carried out and brought the first results confirming the usefulness of this approach. The next steps will be to add semantics to the cultural heritage objects, in order to help the recommendation system to be more relevant in the finding of links between keywords and objects. For that, we plan to use CRM CIDOC as a basis. We also intend to improve the sensorimotor immersive part of the system. Indeed, an enactive system must consider the embodied nature of the user (i.e. movement of eyes, gestures or physiological reactions) to better anticipate his state or intention. It must be also crucial to improve the capabilities of the animations and graphical representations to afford the user's actions and to arouse his perceptions. Different affordances will be studied. Next, we will have to carry out some evaluations of the usability and acceptance of the system. We plane also to study how to evaluate the improvement of the understanding of information through our system. Another interesting cue is the use of this kind of system to construct an interactive authoring tools to populate informed databases.

REFERENCES

- [1] N. Bonnel, A. Cotarmanach, and A. Morin. Meaning Metaphor for Visualizing Search Results. *Proc. of the ninth International Conference on Information Visualization, IV'05*, 2005.
- [2] B. Bonnis, J. Stamos, S. Vosinakis, I. Andreou, and T. Panayiotopoulos. A Platform for Virtual Museums with personalized Content. *Journal of Multimedia Tools and Applications*, 42(2):139–159.
- [3] P. De Loor, K. Manac'h, and J. Tisseau. Enaction-Based Artificial Intelligence: Toward Co-evolution with Humans in the Loop. *Minds and Machines*, 19(3):319–343, oct 2009.

- [4] M. Kaipainen, N. Ravaja, P. Tikka, R. Vuori, R. Pugliese, . Rapino, and T. Takala. Enactive Systems and E_nactive media. Embodied Human-Machine Coupling Beyond Interfaces. *Leonardo*, 2011.
- [5] J. Lacoche, T. Duval, B. Arnaldi, E. Maisel, and J. Royan. A survey of plasticity in 3D user interfaces. *7th Workshop on Software Engineering and Architectures for Realtime Interactive Systems*, 2014.
- [6] H. R. Maturana and F. J. Varela. *Autopoiesis and Cognition: The realization of the living*. Reidel: Boston, 1980.
- [7] M. Merleau-Ponty. *Phénoménologie de la Perception*. Collection "Tel", 1990. éditions g edition, 1945.
- [8] A. Moere and A. Lau. Toward a Methodology for Agent-Based Data Mining and Visualization. *Proc. of the third Australian Conference on Progress in Artificial Life, ACAL 2007, LNBI 4828*, pages 292–304.
- [9] V. Moere, #160, Andrew, Mieusset, K. Hwan, Gross, and Markus. Visualizing abstract information using motion properties of data-driven infographics. *Conference on Visualization and Data Analysis*, 5295:33–44, 2004.
- [10] A. Noë. *Action in Perception*. Cambridge, MA: MIT Press, 2004.
- [11] G. Proctor and C. Winter. Information flocking: Data visualisation in virtual worlds using emergent behaviours. *Lecture notes in computer science*, 1434:168–176, 1998.
- [12] C. Reynolds. Flocks, herds and school : a distributive behavioral model. *Proc. of the 14th annual Conference on Computing Graphics and Interactive Techniques, ACM SIGGRAPH'87*, pages 25–34.
- [13] E. Sklar, C. Jansen, J. Chan, and M. Byrd. Toward a Methodology for Agent-Based Data Mining and Visualization. *Proc. of Agent and Data-mining 2011, ADMI'11*, 2011.
- [14] F. Sparacino. Sto(ry)chastics : a Bayesian Network Architecture for User Modeleing and Computational Storytelling for Interactive Space. *Proc. of UbiComp, the fifth International Conference on Ubiquitous Computing*, pages 139–159.
- [15] J. Stewart, O. Gapenne, and E. Di Paolo. *Enaction: A new paradigm for cognitive science*. MIT Press, 2008.
- [16] F. J. Varela, E. Thompson, and E. Rosch. *The embodied mind: Cognitive science and human experience*. The MIT Press, 1992.
- [17] D. Wijayasekara, O. Linda, and M. Manic. CAVE-SOM : Immersive Visual Data Mining using 3d Self-organizing Maps. *Proc. of International Joint Conference on Neural Networks*, 2011.