Applying Numerical Optimization to Arrangement of Elements in Spatial Interface for Historical Moscow Center Virtual Reconstruction^{*}

¹ Lomonosov Moscow State University, Moscow, Russian Federation, info@vrmsu.ru
² Institute for Archaeology, Russian Academy of Sciences, Moscow, Russian Federation

Abstract. The article describes a novel approach to representing virtual reconstruction of historical cities cultural heritage. As example, the reconstruction of buildings of the historical city center in Moscow was carried out using preserved plans, drawings of buildings, texts and other information. We want to provide users with an opportunity to see the historical territories restored in virtual reality. Everyone should be able to study historical sources for every object, compare them with the result of the virtual reconstruction, see the process of area transformation over time. The usage of virtual reality spatial interface for displaying information about historical objects is proposed. Interface elements should be placed next to interactive objects (such as historical buildings and their parts, landscape sectors, control buttons).

The task of optimizing the automated arrangement of interface elements in the space of virtual reconstruction was considered. Restriction sets for the layout of interface elements have been introduced. Restrictions were obtained from physiological characteristics of human arms and oculomotor apparatus. The tasks of determining the restrictions, the principles that reduce the probability of placing interface elements outside the restrictions was considered. A hand planar movement hypothesis was proved, it allows us to reduce the dimension of the studied system.

Keywords: Virtual reality \cdot Historical reconstruction \cdot Interface \cdot Optimization \cdot GIS \cdot Restrictions \cdot Eye tracking \cdot Arm Movement.

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

^{*}Supported by Lomonosov Moscow State University and by the Russian Foundation for Basic Research grant 18-00-01684 (K) (18-00-01590 and 18-00-01641)

1 Introduction

The problem of documentation and verification of virtual historical reconstructions became relevant in the early 21st century [1].

For a long time, the problem of historical sources publishing that were used in virtual reconstructions was not in the spotlight [2]. Source publication resembled a simple database [3]. This format is good for publishing sources as a separate study, and is almost unlimited in content. Thus a double result is obtained from the project: reconstruction and sources are published as independent results. Sometimes projects are not generally accompanied by sources.

In the last years more and more scientific projects published their reconstruction in the form of images and video source [4]. There are only a few exceptions that not only describe sources, but also publish them [5]. Such projects traditionally use renderers such as Vray, Redshift, Arnold, Keyshot, Lumion. As the result we have videos and rendered images or video without any interaction with sources.

With the development of technology real-time reconstructions have appeared as a lot of reconstructed objects published at Sketchfab platform [6]. Our previous projects included source uploading on a website [7]. The sources were published on the special pages [8]. These sources work like additional part for interactive reconstruction of Strastnoy monastery. Developing ways of representing historical sources, we decided to use virtual reality (VR). The experiment was intended to be conducted at two monastery complexes, Strastnoy and Chudov ones. All models were implemented in virtual environment. A part of this solution is a verification module for historical reconstruction of cultural heritage in VR [9]. It was the first step in real-time interface presentation in VR. For all monasteries we developed a 2D sources verification module (Fig. 1).



Fig. 1. Example of the verification module.

This allowed us to bring the presentation of historical sources to a new level. But we also faced with some limitations. For example, each building reconstructed according to different amounts of sources. In some cases there were 2–3 images, in other more than 10.

During the reconstruction of the Chudov Monastery, an additional task was the integration of sources into decorative elements and interiors, which changed several times during the period of the monastery's existence while maintaining the basic forms of objects. This once again complicated the structure of source presentation and made it difficult for a user to interact with an interface.

The next step in improving the interactive capabilities of the user to assess the information potential of the sources used, to verify 3D models is to immerse the user in a virtual environment. Integration of reconstruction sources into VR requires new algorithms for creating a three-dimensional 3D interface. When developing such interfaces, it is necessary to take into account a large number of parameters of human movement in virtual environment, which will be discussed below. An additional condition for us was the historical landscape, the reconstruction of which is one of the main objectives of our project. 3D models of historical buildings are placed on the landscape, in compliance with the scale of objects and features of the relief. This allows the user, moving in a reconstructed historical urban space, to expand interaction with both 3D models and sources used for building those models.



Fig. 2. Example of a landscape reconstruction of Moscow Belyi Gorod area.

Historical part of our actual research project is aimed at the virtual reconstruction of Moscow historical center (Belyi Gorod) landscape and historical buildings located on its territory. To reconstruct the landscape of Belyi Gorod and dominant historical buildings in 16th – 18th cc. we used archaeological and geological measurements of the Belyi Gorod relief, preserved plans and drawings of historical buildings, old photos, textual sources and other historical materials. The results include both sources and reconstructed objects data.

2 Numerical optimization of the interface element layout

2.1 The task of the interface layout numerical optimization

Each object of the virtual reconstruction is provided with reference information and with interactive elements that allow to open the Menu, to change the object state or to get access to a historical source. The task is to locate information interface elements and interactive interface elements. Despite some differences between information and interactive elements, this article suggests a general approach to the search of the proper position for both types of elements.

To determine the location of interface elements on a computer screen, the Fitts law [10] and its various extensions are often used. This law describes an empirically determined relationship between the duration of motion to the target in the plane, the distance to the target and its size. It works well for planar interfaces, but virtual reality leads us to a three-dimensional problem.

A user interacts with a 3D interface using his eyes and his own body motion. The simplest way — user places his hand in a specific point in a virtual space. Obviously, elements should be placed inside a limited area where it will be convenient for a user to interact with them. They shouldn't overlap with other objects of the virtual scene and with each other. Here, we describe ideas how to define a priority zone for the interface elements location, taking into account existing restrictions.

Interface elements should be considered as closed sets, further they will be represented by a center point. Let there be a dynamic system that describes the movement of an eye or hand: $\dot{x} = f(t, x, v)$, where $v \in V$ is a perturbation vector. Perturbations may be inaccuracies in determining parameters of the hand, the position of the target interface element, and so on.

We assume that the movement goes from an initial state to an interface element or from one element to another. In the latter case we take the the first element position as initial conditions, and the position of the other as terminal. We define the perturbation vector. Then the problem of optimizing the positions of interface elements can be posed.

We enumerate all the elements of the interface and define the probabilities of the transition from the *i*-th to the *j*-th element as the coefficients $k_{x_ix_j}$ of the weight of each transition from x_i -position to x_j -position. The probabilities are defined according to the information or functionality attached to the interface element.

As was said before, elements should not overlap and intersect with other objects in the scene, which imposes restrictions on the set of feasible system solutions. At the same time, it should be possible for a user to reach all of the interface elements, both information (by eyesight) and interactive (by eyesight and hand). All these restrictions can be defined as Υ , a set of the interface element position restrictions, $x \in \Upsilon$.

As a result, if perturbation vector v is given, we have the problem of minimizing the weighed sum for N interface elements by placing them in the restriction

set
$$\Upsilon$$
:

$$J_N = \sum_{i,j} k_{x_i x_j} J(x_i, x_j, v) \to \min_{x_m \in \Upsilon}, \qquad i = 1, 2, \dots N, \quad j = 1, 2, \dots N, \quad (1)$$

where $J(x_i, x_j, v)$ is an optimal movement time from the *i*-th to the *j*-th element or more complex functional, for example, energy expended for the transition. It can be determined from a model or from an experiment.

The optimal positions of interface elements affected by perturbations can be found using game theory. Let the lengths of the joints, masses, etc. be disturbed. The relative position of the interface elements will be our control.

We can get the antagonistic game Γ : the player in charge of controls minimizes the functional, the player in charge of perturbations maximizes it. The lower estimate of the quality of the interface is the following value:

$$\min_{x_m \in \Upsilon} \max_{v \in V} \sum_{i,j} k_{x_i x_j} J(x_i, x_j, v).$$
(2)

This formulation of the problem allows us to optimize the layout of the interface elements, taking into account "hard" restriction set Υ for the users hands and eye movements. But in a real situation, some locations of objects lies on the Υ set boundaries. Although they remain reachable, can cause discomfort for a user when looking at them or when trying to reach them. It is required to find a set of comfortable arrangement of interface elements. Let us call this set Ξ "soft" restrictions. We can define a penalty factor $s_m(\rho(x_m, \Xi))$ when placing elements outside this set, where ρ is a distance between an interface element x_m and the set Ξ . If $x_m \in \Xi$, then $s_m = 1$.

Taking into account the penalty factor, the problem (2) could be formulated as follows:

$$\min_{x_m \in \Upsilon} \max_{v \in V} \sum_{i,j} k_{x_i x_j} s_i s_j J(x_i, x_j, v) \tag{3}$$

To solve this problem, dynamic programming methods [11] are applied. Using them we could automatically optimize interface element placement for every object of the historical reconstruction.

The purpose of this article is to define "soft" restrictions such as human hand movement restrictions and visual restrictions due to possible intersensoral conflict.

We also summarize results of the eye and arm movements analysis. From eye movements we find the criteria for "soft" restrictions on elements placement and size. For arm we describe a hypothesis of planar movement which allows us to reduce the dimension of the problem for goal-directed hand movement from one interface element to another.

2.2 Arm motion analysis

Human arm mobility Let us describe the restrictions imposed on the optimal positions set of interface elements that arise due to the limited reach of a persons

arm. We define the reachability set of the end arm effector, i.e. the set of all permissible positions of the hand. For this, we need parameters and possible rotations of each link in the arm.

Many possible arm positions are constructed and they describe the set of constraints for interactive interface elements. We used a book [12] describes in detail the results of the parameter measurements of the human body parts. More than 200 men and women from the space crew and NASA employees were invited as subjects. These personnel are in good health, fully adult in physical development, and an average age of 40 years. Body masses were measured as well as masses of body parts (neck, head, shoulder, forearm, hand, etc.), their lengths and centers of mass positions, inertia moments, volumes, changes in the centers of mass in various poses, permissible movements for parts of human bodies. Reachability sets were found for various segments.

From [12] the average values of the parameters of the right hand for men were taken and used for calculations: l_1, l_2 are the lengths of the shoulder and forearm. According to [12] $l_1 = 0.366$ m., $l_2 = 0.305$ m.

To build the reachability set for the hand joints, we used data from [12]. The table 1 below shows the limits of feasible rotation angles for different axes for most male subjects. To simplify the problem, we suppose that a human arm rotates only around its longitudinal axis. This statement will not affect the reachability set, therefore it will be built only for the shoulder and forearm.

Description of rotation		Boundary angles values			
		rad.			
Horizontal abduction (Fig. 3, 4)	188,7	3,292			
Shoulder lateral rotation (Fig. 3, 5A)	96,7	1,686			
Shoulder medial rotation (Fig. 3, 5B)	126, 6	2,208			
Shoulder flexion (Fig. 3, 6A)	210,9	$3,\!679$			
Shoulder extension (Fig. 3, 6B)	83,3	1,453			
Elbow flexion (Fig. 3, 7)	159,0	2,773			

Table 1. Various joints rotation feasible angles (NASA, [12]).

In Fig. 3 we can see the measured angles and angles of reference. All figures have angles which are marked by letters A and B. Not all of these angles are indicated in Table 1, for example, the angle B of elbow flexion/extension is indicated for the horizontal direction, since the sum of the angles A and B is 180°. Also, the sum of the angles A and B for flexion and extension of the forearm is 180°.

When we change the angle of the shoulder, the feasible angles of rotation of the forearm also change. This statement was also described in [12]. Fig. 4 is an example of a reachability set presented for a forearm with the restrictions specified in the Table 1.

Reachability sets for the whole arm are obtained by combining reachability sets for the forearm for all valid shoulder locations. We use the human arm parameters values from [12] as average data. A tracking system of a virtual reality device



Fig. 3. Rotation of human arm: shoulder and forearm (NASA, [12]).



Fig. 4. Reachability set for a forearm.

allows us to specify it for every user. It helps us to determine "soft" restriction set Ξ for interactive interface element placement, but we do not take into account user's comfort when moving hands to extreme positions.

Plane motion hypothesis According to the problem (3), the definition of the functional J and the penalty factors s_m (describing user's discomfort during interaction with elements that are not in the set Ξ) could lead us to very different interface element placing.

One of the ways to define J and s_m is an energy criteria, when we consider the equations of motion for hand and calculate energy consumption for the transfer from one interface element to another (especially for extreme positions). The spatial problem of moving an arm from one position to another is complex and has an infinite number of solutions even if a wrist trajectory is pre-defined.

That is why a planar problem was considered. We determined some situations when the spatial movement of an arm can be considered planar. A hypothesis test was carried out: the movement of the hand will be planar if there is a plane in which the hand lies at the initial and final moments. This statement was proved experimentally.

The experiment involved 17 right-handed students (4 females, 13 males). They were asked to hit the target as quickly and accurately as possible with a pointer (Fig. 5), located on the board in front of them. Targets (goals) are square holes in the plane with the sides of 3, 4, 5, 6 cm. The order of achieving the goals

is predetermined. The initial position of the arm, from which the movement started, was also set and marked. A mark for the palm was put on the pointer; during the experiment it was not allowed to change the position of the pointer in the hand. After each movement, the subject's arm returned to its original position and remained there for 3–5 seconds. To record the experimental data, markers of the video analysis system were fixed on the shoulder, forearm of the human right hand and pointer, to track their positions in space. We recorded the coordinates and orientation of the specified markers and time.



Fig. 5. Target and pointer with the object of the video analysis system.



Fig. 6. The pointer's trajectories compressed along the axis of motion and rotation Oy. The trajectories are rotated around the axis Oz and shifted to match the end points.

The data obtained were interpreted using the MathWorks MATLAB software. For the trajectories of each marker, the approximate planes to which they belong were constructed, the average trajectory deviations from the corresponding planes. Then we calculated the values of standard deviation for the previous deviations. The angles between the planes of the motion of the tracking system markers were also found.

Tables 2 and 3 show the average values of the obtained values for all movements.

Table 2. Deviations of the trajectories of markers from the plane of motion.

	Average (mm):	Standard deviation (mm):
Shoulder	1,89	1,85
Forearm	1,89	1,84
Pointer	$5,\!6$	4,84

Table 3. Angles between the planes of motion of markers.

	Average	(deg.)	Star	dard o	leviatio	on (mi	m):	
Shoulder-forearm	0,4710	0,4575						
Forearm–pointer	1,0633	0,4901						
movements count	60 50 - 40 - 30 - 20 - 10 - 0 0	5	10	15 deviatio	20 20	25	30	35

Fig. 7. Average value distribution for the trajectory deviation from the plane of the pointer movement.

Fig. 7 shows the average values of the distribution calculated for the trajectory deviation from the plane of pointer movement. The average deviations from the plane of movement and their standard deviations were 2–3 times greater for targets not lying in the plane of the initial arm position. The average range of motion (the distance between the start and end points of the trajectories) is about 400 mm. It is two orders of magnitude greater than deviations from the plane. The angles between the planes of marker motion are small (see Table 3).

So the movement trajectories can be considered planar. This allows us to use planar models of arm motion dynamics [13] to describe transitions between interface elements.

2.3 Eye motion analysis

The eye motions are quite complex. The motion from point to point is only one of tasks solved in this field by human nerves system. How we see, the convergence angle is changing during saccade, but this is not the only situation when it's happen. When we examined the interface element our eyes are permanently moving. Eye motion control mechanisms provide clear vision [14]. Any human movement requires such resources as energy, ions. There are a lot of adaptation mechanisms in human organism, but sometimes our organism use to spend more

resources to solve really important problem. Such task is a visual perception of the world.

Following Filin our eye make two or more small saccades per second when gaze is stable [14]. That small saccades are bit different in amplitude and latency. We suppose that investigation such movements could be criteria as an index of comfortable vision. It could help us to find the "soft" restrictions set Ξ for informational interface elements and to define the penalty factor.

Eye movement analysis in daily life conditions is not simple problem. There are a lot of disturbance because of person head rotations, walking, breathing, eye tracking system displacement. The first high accuracy eye tracking systems were huge, hard fixed, and persons head fixed hardly too. In such conditions there are a lot of research on the functioning of the visual system were produced. Now it is necessary to shift at technics we have in new conditions when eye tracking system is placed as glasses, there are some disturbances and we need real time information. The probabilistic moments of eye tracker signals can be a characteristic of eye movement strategy.

Probability moments in eye movement analysis In ideal world we can identify saccade parameters and interpreted eye movement as physiological answer on presented to person situation. Usually we have a number of artefacts in eye tracking data. I this situation we offer to use the probability moments to class the visual surround.

In the work [15] were shown there is a comfortable distance for examine the object. We prolong their investigation for object in virtual reality. In virtual reality we have a conflict between accommodation and convergence because a distance to the screen is fixed.

We make an experiment using Panoramic virtual reality system, Arrington research eye tracker and the accompanying software. The process of the experiment is as follows. After a participant puts on stereo glasses and assumes a comfortable sitting pose, a calibration is performed using built-in tools. After confirming successful calibration, the main testing application is launched.

Participant is then shown yellow spheres on a black background spawning in front of him at distances from 0.4 to 6 meters in random order. Angular sphere size was constant value and equals 0.7 degree. Each sphere is displayed for 60 seconds, then it changes color giving the participant a cue for blinking several times. This behaviour triggers changing distance to a sphere. After the last sphere is shown, only black background is shown and we let the participant to take a rest. Then we change the participant's distance to the screen from 0.4 to 6 meters in random order six times.

At the end we let the participant to remove the stereo glasses and finish his participation. A total of 12 persons took part in this study. All of them gave informed consent to participate.

Usually three component of eye position is registered and analysed. We consider the angle between visual axis of the eyes. For this quantity we analysed probability second moment [16]. Values are presented in the Table 4.

Title Suppressed Due to Excessive Length 11

Screen dist.	Sphere distance					
	0.4 m	$0.68 \mathrm{~m}$	$1.18 \mathrm{~m}$	$2.03 \mathrm{~m}$	$3.49 \mathrm{~m}$	6 m
0.4 m	0.0224	0.0188	0.0075	0.0197	0.0081	0.0078
0.68 m	0.0163	0.0160	0.0064	0.0093	0.0070	0.0084
1.18 m	0.0116	0.0086	0.0058	0.0081	0.0108	0.0135
2.03 m	0.0292	0.0093	0.0127	0.0116	0.0062	0.0080
3.49	0.0212	0.0107	0.0078	0.0098	0.0067	0.0130
6 m	0.0163	0.0093	0.0085	0.0106	0.0074	0.0098

Table 4. The STD mean value in radians [16].

We find that probability second moments correlated with degree of mismatch in real and imagine distance. In most comfortable situation the dispersion value from eye position were lowest.

In a comfortable situation, a person has saccades in one direction and vergent. In a mismatch situation as sensory conflict the proportion of vergent saccades is increasing. This phenomena we see in our investigations. In study [15] oculogram probability moments had not significant divisions. But in 3D case we see significant divisions correlated with mismatch between real and virtual distance.

This way we can make a criteria for "soft" interface restrictions Ξ : the time when the convergence angle STD is high need to be minimal. According to the STD changing, we can define the penalty factor $s_m(\rho(x_m, \Xi))$.

3 Conclusions

The article describes a novel approach to the interface design for virtual historical reconstruction. The problem of optimizing the location of spatial interface elements is stated. The solutions of some related problems are given.

Human arm parameter analysis gives us the restriction set for the interactive elements. The hypothesis of targeted hand movement planarity is confirmed.

According to the eye motion analysis, informational interface elements with which long interaction is supposed need to be placed a meter or more away. The time of interaction with close placed object must be minimized. The next stage is to apply frequency analysis which allows us to slightly change the interface in real time. During the interaction with the interface, cutoff frequency for each user can be specified and virtual environment adapts to it.

This study allows us to create a convenient virtual reality interface giving information about reconstructed objects.

References

 Kuroczyski, P., Hauck, O., Dworak, D.: 3D models on triple paths - New pathways for documenting and visualizing virtual reconstructions. In: Mnster, S., Pfarr-Harfst, M., Kuroczyski, P, Ioannides, M. (Eds.), 3D Research Challenges in Cultural Heritage II. Springer LNCS, Cham, pp. 149-172. (2016)

- 12 L. Borodkin et al.
- 2. Abrau Antiqua (in russian), http://abrau-antiqua.ru. Last accessed 30 Jan 2020
- 3. JCB Archive of Early American Images, https://jcb.lunaimaging.com/luna/servlet. Last accessed 30 Jan 2020
- 4. Paris 3D: Through the Ages Dassault Systmes, https://youtu.be/-64kHmCJGMA. Last accessed 30 Jan 2020
- Dworak, D., Kuroczynski, P.: Virtual Reconstruction 3.0: New Approach of Webbased Visualisation and Documentation of Lost Cultural Heritage. In: EuroMed 2016: Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, pp. 292–306. (2016)
- 6. Heritage 3D models, https://sketchfab.com/tags/heritage. Last accessed 30 Jan 2020
- Borodkin, L., Valetov, T., Zherebyat'yev, D., Mironenko, M., Moor, V.: Reprezentaciya i vizualizatciya v onlaine rezul'tatov virtual'noy rekonstrukcii. In: Istpricheskaya Informatika, no. 3–4, pp. 3–18. (2015). (in russian)
- 8. Sources for building a virtual reconstruction (in russian), http://www.hist.msu.ru/Strastnoy/Source. Last accessed 30 Jan 2020
- Borodkin, L., Mironenko, M., Chertopolokhov, V., Belousova, M., Khlopikov, V.: Technologii virtual'noy i dopolnennoy real'nosti (VR/AR) v zadachah reconstrukcii istoricheskoy gorodskoy zastroiki (na primere moskovskogo Strastnogo monastyrya). In: Istoricheskaya Informatika, no. 3, pp. 76–88. (2018). https://doi.org/10.7256/2585-7797.2018.3.27549 (in russian)
- Fitts, P.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47 (6), pp. 381–391. (1954)
- Bellman, R.: Functional Equations in the theory of dynamic programming. Proc Natl Acad Sci USA. 1955 Jul 15; 41(7), pp. 482–485. (1957)
- NASA Man-Systems Integration Standards, Revision B, July 1995, https://msis.jsc.nasa.gov/. Last accessed 30 Jan 2020
- Bonilla, F., Lukyanov, E., Litvin, A., Deplov, D.: MATHEMATICAL MODEL-ING OF THE UPPER LIMB MOTION DYNAMICS. Journal Modern problems of science and education, part 1. (2015)
- 14. Filin, V.: Avtomatiya sakkad. Moscow: Publishing House of Moscow State University. (2002) (in russian)
- Kaspransky, R., Muratova, E., Yakushev, A.: The use of videooculography to assess a comfortable distance to the target. In: Biomechanics of the eye / Ed. Iomdina E.N., Kositsa I.N., pp. 166 – 168, MNIIGB them. Helmholtz. (2005) (in russian)
- 16. Kruchinina, A., Chertopolokhov, V., Yakushev, A.: Metodika polucheniya chislennogo kriteriya nalichiya sensornogo konflikta na etape sozdaniya vizual'nogo kontenta. In: Zapis' i vosproizvedenie ob"emnyh izobrazhenij v kinematografe i drugih oblastyah: IX Mezhdunarodnaya nauchno-prakticheskaya konferenciya, pp. 260–267. (2017) (in russian)