Data representation, transparency and colour improvements for AR4Morduc client

Alessandro Caniglia

UH, 15 March 2010 – 12 April 2010

Version: 3.0

Change log:

('+' for added, '-' for deleted, '*' for modified content)

- * diagram 1
- * image 3 caption
- * chapter 2 introduction
- * paragraph 2.2 (and sub-paragraphs)
- * chapter 3
- + chapter 4

Index

1 Introduction

Livatino et al., in [1], proposes "*a method to simultaneously and coherently present visual and laser sensors information through an augmented reality visualization interface further enhanced by stereoscopic viewing. The use of graphical objects is proposed to represent proximity measurements which are superimposed and suitably aligned to video information through image processing. This new methodology enables an operator to quickly comprehend scene layout and dynamics, and to respond in an accurate and timely manner. Therefore the human-robot interaction is expected to be intuitive, accurate and fast. The use of graphical elements to assist tele-operation, sometime discussed in the literature, is here proposed following a systematic approach and developed based on authors' previous works on stereoscopic tele-operation*". The approach is experimented on a real tele-robotic system where a user operates a robot (called 3Morduc[1] and also known as Morduc) located approximately 3,000 kilometers apart.

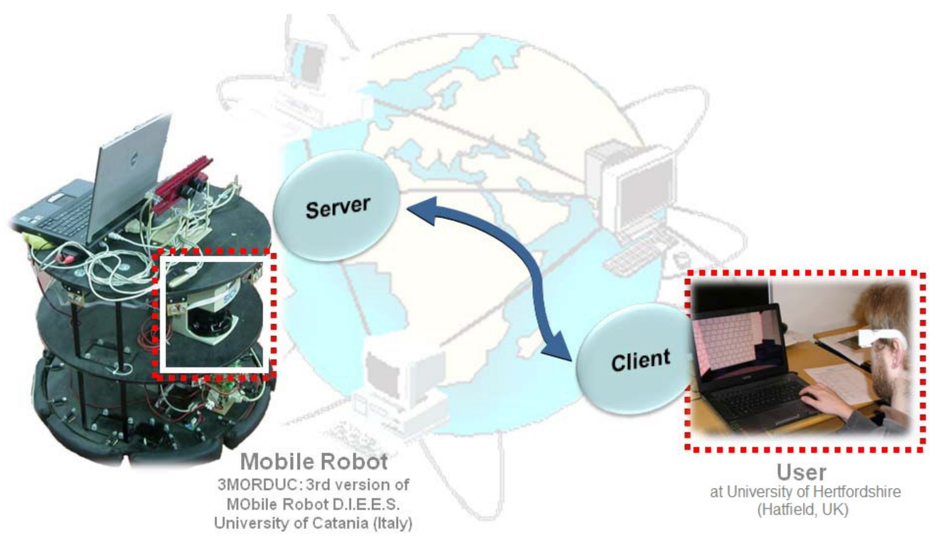


Image 1: Tele-robotic system used to develop 3Morduc clients

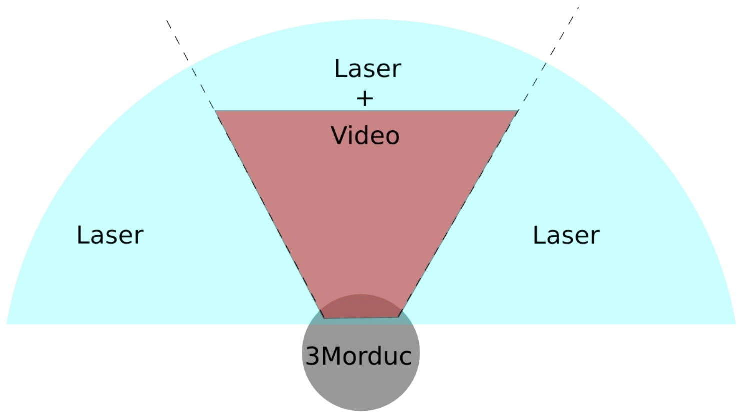


Image 2: Currently 3Morduc used sensors

Elder version of our tele-guide client have been developed following other approaches; more details about them are in [12, 13, 14].

The current version of the client is called AR4Morduc[2] due to implemented Augmented Reality functionalities.

In the current version of AR4Morduc client, image processing and computer graphics are used to add information about distances to camera grabbed images of the environment surrounding Morduc mobile robot.

A brief discussion about this client version and some images of Augmented Reality offered by AR4Morduc software can be found in [1]. AR4Morduc is the first client developed by our research team which uses both camera frames and laser information to create an Augmented Reality visualization of the world surrounding 3Morduc. This approach presents distance information in

1 3Morduc stands for 3rd version of MObile Robot D.I.E.E.S. University of Catania (Italy)

2 AR4Morduc stays for Augmented Reality for 3Morduc

different modalities:

- Reality mode (only camera frames);
- AR composed by camera frames + coloured areas;
- AR composed by camera frames + coloured segments + written distances;
- AR composed by camera frames + coloured areas + coloured segments + written distances.

Typically, during AR4Morduc tele-guide, user can see an AR environment composed by a real background and computer graphics foreground (except in Reality mode).

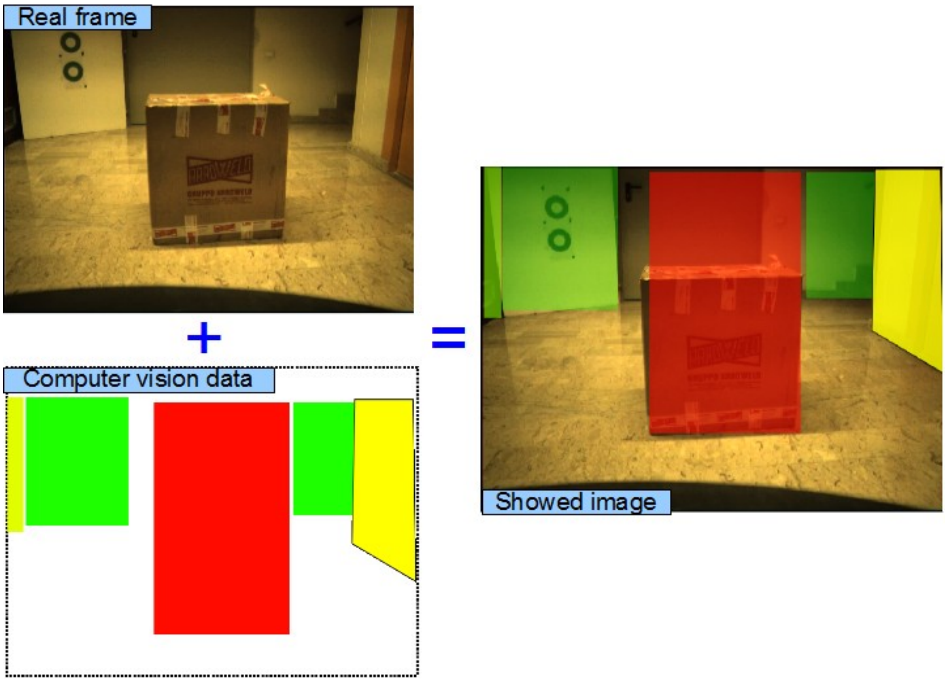


Image 3: Process of composition for a showed image in AR4Morduc: showed image has the real frame as background and computer graphic data as foreground

We use stereoscopic visualization during development and tests of AR4Morduc client because [12] says that "*it is possible to find works demonstrating that stereoscopic visualization may provide a user with a higher sense of presence in remote environments because of higher depth perception, leading to higher comprehension of distance, as well as aspects related to it, e.g. ambient layout, obstacle perception, and the accuracy of the manoeuvres*". In fact, from the same paper, "*the 2-D display systems commonly used in robot tele-operation suffer from many limitations. Among them are misjudgement of self-motion and spatial localization, limited comprehension of remote ambient layout, object size and shape, etc. The above limitations lead to unwanted collisions during navigation, as well as long training periods for an operator*".

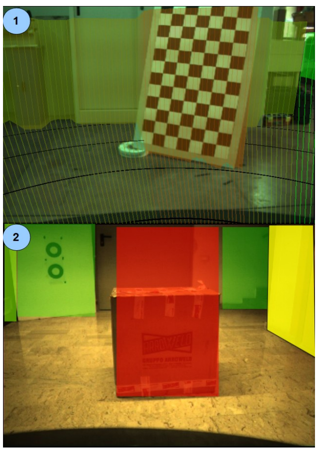


Image 4: Examples of OpenGL (1) and OpenCV (2) AR images

Livatino et al., in [12], test results did not show significant improvement when estimating self-motion and collision avoidance using stereoscopy systems related to 2D systems. To cope with these shortcomings, we are concentrating our research on AR. Furthermore, the use of stereoscopic visualization in addition to AR may represent a further improvement to the proposed approach.

Current version of AR4Morduc is based on computer vision processing; this allows us to represent additional layers on the original video image. Our current works are focused on the usage of computer vision and computer graphics. In particular we use OpenCV[3] and OpenGL libraries. We think that OpenGL will help us to show foreground data in a more reliable way for user tele-guide experience. Moreover, OpenGL SDK offers easier tools for solid object creation, such as solid icons and 3D object rendering, special effects (e.g. shadow), change the point of view of the scene dynamically and the possibility to use stereoscopic systems with ease.

Other interesting information about our research activity (about 3Morduc and its AR4Morduc client) can be found in [1].

1.1 Target of the following discussion

An improvement has been planned for Morduc client: development of a better representation of computer graphics data (distance measurement, say) using transparencies and colours more properly.

3 OpenCV (Open Source Computer Vision) is a library of programming functions for real time computer vision.

2 Distance representation

Currently, we can use camera frames (they could be photographs or video frames as well), distance data measured by 2D laser scanner and odometer data.

About foreground computer graphic objects, three fundamental colours are used to suggest distances: green, yellow and red. Each coloured area is painted with a color chosen on distance information retrieved by the 2D laser scanner. The client generate a color look-up table, where the color "red" represents the smaller range value (closest distance) and the color green represents the biggest range value (farer distance) [1].

Our target is to organize distance information better on camera frames to create an effective experience of the environment (real and graphical) for users.

Two aspects will be discussed in this first section: a better usage of transparency and a better usage of colours. A third topic is a different data representation using different shapes (apart from simple coloured area, coloured segments a0nd written distances) and data representations which should distract less user tele-guide activity.

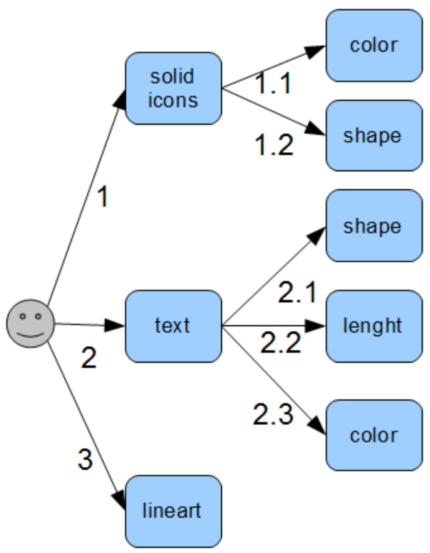


Diagram 1: Users' decisional tree about data representation over a background using transparency

In [2], a study, about transparent user interfaces and information content representation, is explained.

Our typical test environment is a piece of real world (a room with boxes on the floor). The author of [2] assures that we are in good conditions to use transparency for foreground graphic objects (real world as background and semi-transparent computer graphic objects as foreground).

We think that qualitative results of [2] could be interesting: users use to follow the decisional tree showed in diagram 1 about level of attention related to background and foreground objects. In particular, test results show that users understand and perceive sooner what solid icons communicate than simple text and line arts, which are the worst way to communicate a meaning using foreground transparency over a background image.

Also in [3] usage of icons is suggested according to some Gestalt laws: people use to understand quickly well known shapes (coloured and not) because they associated them default messages such as danger, information, direction, and so on. We plan to use transparency, colors and other shape instead of (or with) coloured areas and segments, inside AR environment, to make more understandable and less confusing the showed data. In fact, we suppose that a coloured area (with the same size of the entire object) could cover some interesting particulars of the object itself; a way to avoid this it is to increase the percentage of transparency of the graphic layer, but in this way

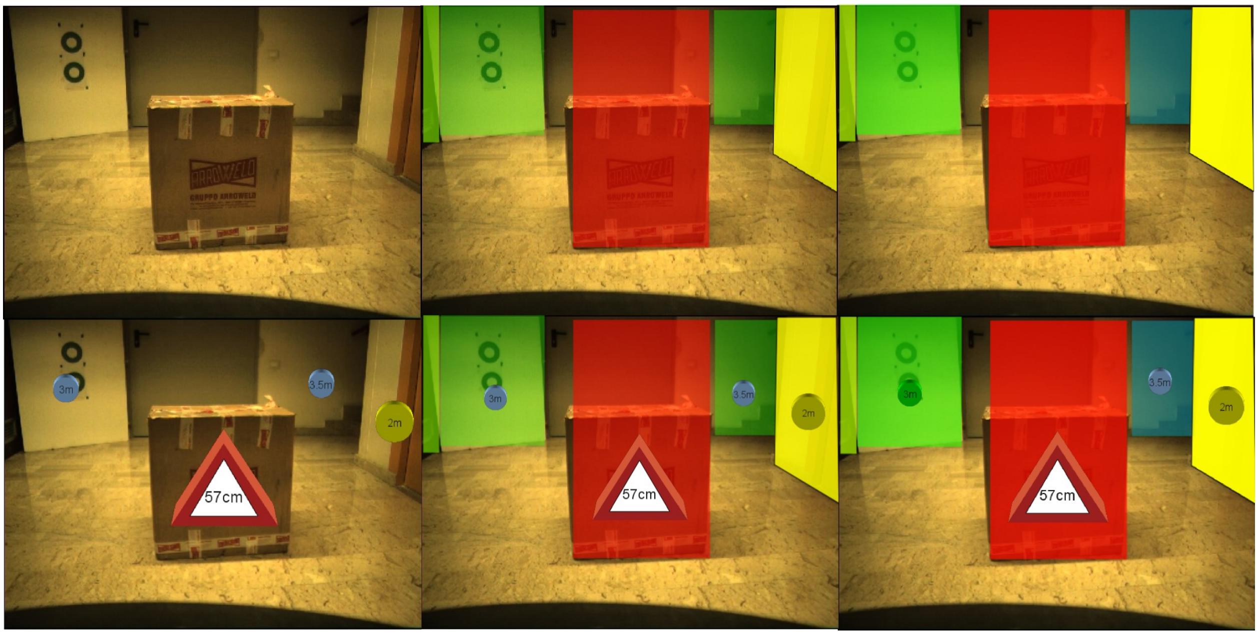


Image 5: Comparison between different data representation. First row: only real frame, color stripes, color stripes with blue color; second row: solid icons, solid icons plus color stripes, solid icons plus colored stripes with blue color

graphic objects could become too transparent and so less visible to the user. Currently, it is not possible to change the transparency value during user tele-guide activity (it is set before the tele-guide starts) and if this would be possible it can not be easy either. A different presentation of graphical object could be a better choice.

For example we could think about the following key points trying to explore new approaches for our AR application:

- dynamic settings (e.g. dynamic/adaptive transparency)
- shapes (e.g. solid icons, different shapes according to the message to be understood by user)
- colours
- special effects (e.g. blinking, fading, animations)
- blending elements (e.g. color, transparency, monocular depth-cues)
- additional sense experiences (e.g. using sound, smell, taste, touch)
- stereoscopy (e.g. perspective)

In the following paragraphs we suggest complementary approaches which could be used to communicate distance information to tele-operator.

2.1 Dynamic and adaptive settings

As you have read above, the current version of AR4Morduc let user decide settings at the beginning of tele-guide or during it (but only some settings can be changed). We think that the possibility of changing settings on-the-fly is useful because user visualization can be adapted to the particular environment or activity the robot is going to perform.

But an annoying fact could arise: in particular circumstances, user needs to change settings many times during a single tele-guide session due to lighting changes, for example. In this case a dynamic, or adaptive, behaviour of the client could make user experience more effective.

We can change and manage many settings during a tele-guide session, e.g. brightness, focus/sharpen, contrast, point of view perspective.

We can offer user the choice to modify these settings at run-time, i.e. if he needs particular visualizations, and, at the same time, the client tries to create the best possible visualization using adaptive functionalities such as contrast stretching, color balancing, gamma transformation, histogram equalization for each showed frame.

2.2 Shapes and colours

MacDonald in [21], says that "*for effective design of presentations on computer displays, use color in conjunction with the other visual variables of position (x,y coordinates), value (lightness), size, shape, orientation and texture*". In this paragraph we cover different aspects of data representation using colours, icons, shapes and other ideas.

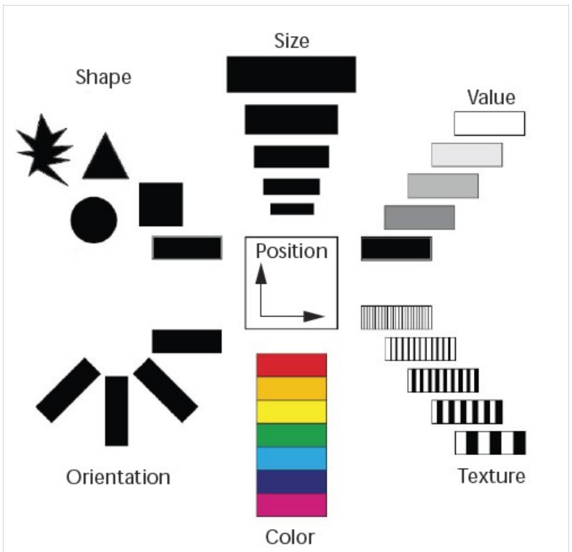


Image 6: Visual variables [21]

2.2.1 Icons

Solid icons (in front of the object actually covered by a coloured area) could be used to show distance information (see image 4 for examples). MacDonald in [21] suggests that "*the general rule is to make form indicative of function*", so symbolic signs could be used to exploit user knowledge about conventional mapping between signifier and signified. It could be chosen an arbitrary mapping, using new symbols too, but in this way the relationship must be learned [8].

In [7] is discussed the usage of circle shapes to inform user about distance in a sort of navigation software. Another idea is to use different shapes according to the message to be communicated, e.g. well known road signs or coloured solid shapes superimposed on coloured rectangular areas.

2.2.2 Colours

We use to associate a particular colour to each object (detected by 2D laser scan) of the background according to its distance from 3Morduc. It is important to create an effective correspondence between each distance and the particular colour. MacDonald in [21] suggests that "*the principle of of the least astonishment" (that guides dialogue design) applies equally to color. The user should quickly recognize each functional element and not get confused by unexpected changes in the assignment of different colors to similar functions*". Secondly, colours can provide a very effective means of increasing the information content of a display or making it easier to interpret [21].

About the usage of colours to suggest information about distance, Williams in [10] confirm our colour choice: red, yellow and green are good colour hues because they can be viewed together without refocusing and they represent respectively danger, caution, safety in Western European culture (and also in American, Japanese and Chinese one). Secondly "*bright red, orange and yellow are reserved for cases of emergency and conflict alert*" [21].

We could think to add blue hue to this set of colours for very far objects; indeed, blue suggest people calm (*it is the colour of the sea and the sky and hence it has associations of depth and stability* [21]).

Colour	Associations
Red	danger, aggression
Green	safety, environment
Yellow	Hazard, illness
Blue	depth, peace, stability

Table 1: Colour associations

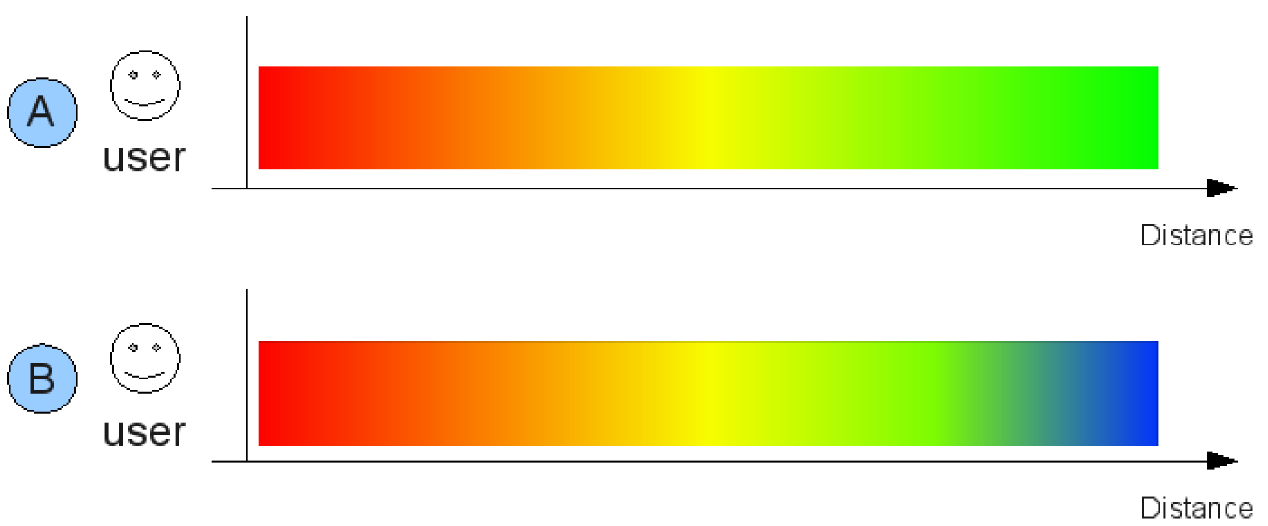


Image 7: Usage of different colours to suggest distance information: approach A uses 3 colours (red, yellow, green); approach B uses 4 colors (red, yellow, green, blue)

According to MacDonald [21] and Hunt [20], human eye's sensitivity to blue is very much lower than to red and green so blue is a good candidate to represent far objects (which are considered in our research less important than close objects); in fact, "*researchers estimate the relative abundance of the three cone types to be in the ratio red:green:blue = 40:20:1*"[21]. This ratio suggest us that our choice should be correct: red colour is a suitable for the most dangerous (close, say) objects because human eyes have more sensitivity to it than other used colours (green, yellow, blue).

Furthermore, according to MacDonald, human eyes change their convexity according to focused colours: the lens become more convex to focus on red colours; to focus on blue colors, the lens become less convex. "*These factors contribute to the apparent color-depth effect called **chromostereopsis**: for the majority of observers, red appears to advance while blue appears to recede*" [23]. MacDonald, in [21], suggest to avoid juxtaposing strong red and blue to prevent unwanted depth effects on display, but we could just exploit this effect to create a depth effect for distance object information.

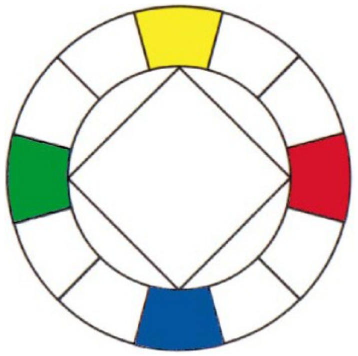


Image 8: Tedrad-opponent hues

According to MacDonald, we could create harmony in our client design using a limited group of colours (a colour palette, say) that look pleasing in combination. In particular, as explained above, we think to use 4 fundamental **tetrad-opponent** colours which have a specified relationship in the 2D hue circle.

About colour selection, MacDonald suggests to use bright, highly saturated colours to grab user attention [21] but not for long periods because user's eyes could be peeved.

2.2.3 Coloured areas

Currently, AR4Morduc displays distance information using coloured computer graphics areas covering video image objects. MacDonald, in [21], explains that “*suppression of the eye's scanning movements can fixate a strongly coloured image on the retina; this temporarily reduces the sensitivity of the photoreceptor, leading to after-images*”. Murch says that “*it is unwise to use large areas of bright color in display”* [22]. For these reasons, we think to change the representation of distance information using different approaches to avoid the after-image effect and to reduce user stress caused by shown bright colour areas on the screen for a long time. A possible solution is to use smaller coloured area instead of full-size ones.

2.2.4 Gestalt laws

According to [21, 24], Gestalt laws are good guidelines for display composition. Zakia in [25], says that “*laws of proximity, similarity, continuity, closure and figure-ground provide powerful organizing principles that apply to all aspects of design, including color*”. MacDonald in [21] confirm our idea: “*you can use color for both association, indicating that certain elements in a design have common properties, and for differentiation, indicating that certain elements differ in their properties from the others*”.

2.3 Using more human senses

In our client, currently, we present information to user through images (AR images, say).

It is interesting the possibility of other user senses such as sense of hearing, touch, taste, smell.

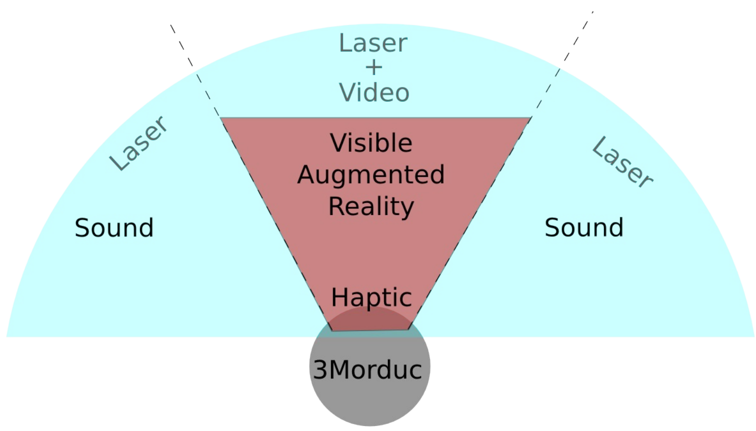


Image 9: 3Morduc used sensors (laser and video) and possible device improvements to communicate information to tele-operator using different human senses

Image 10: Examples of devices to use other human senses

We can, for example, play a particular sound in the left user headphone if an invisible[4] object is very close to the robot; in this way tele-operator could react even if he is not able to see the obstacle on robot left side (turning left the robot could be dangerous in small spaces). This particular application of sound it could be implemented with ease in our system because we use a 180° laser scan which assures a range wider than robot cameras range of view.

About sense of touch, we can exploit different devices which could help user to move the robot in a more natural way (e.g. moving an hand in the direction he would move the robot). Furthermore particular peripherals, e.g. cyber gloves, could be used for tele-manipulation purposes too. Haptic devices could communicate user collisions or very close objects or current movements of the robot (in this way user is able to know if robot is still moving or not).

In future, we could plan to use more sensors to detect objects around 3Morduc and these will give us more data to exploit. Probably, other devices could be used with a growing number of sensors because only visualization could become too crowded and confusing for tele-operator.

2.4 Depth cues and stereoscopy

User should have a realistic vision of the world surrounding the robot. To achieve this effect we can use different approaches, i.e. monocular depth-cues, stereoscopy.

Neri says that "*the perspective is the strongest depth cue. It is based on the relationship between foreground and background objects. When the perspective is present we feel as all the parallel lines converge on a point, the vanishing point, situated on the horizon line. The perspective deforms the object shape, in fact, the closer parts become bigger and the further parts become smaller then the objects closer to the horizon seem more distant. Similar to the perspective depth cue is the relative size. It is based on the knowledge that closer object are bigger than those which are in the second plan"* [15].

We are able to create a more realistic presence of operator inside the robot scene thanks to AR and stereoscopy. Thanks to Graphic libraries, i.e. OpenGL, we can create foreground objects following monocular depth-cue principles in addition to stereoscopy vision. In this way tele-operator is lead to think and behave as he is actually in the same place of the Morduc (tele-presence[5], say).

4 Invisible in the sense the object is not inside robot range of view (it is not framed by robot cameras)

5 Telepresence refers to a set of technologies which allow a person to feel as if they were present, to give the appearance that they were present, or to have an effect, at a location other than their true location.[16]

3 More about data representation and usage of transmission delay period

Humans have remarkable perceptual abilities, that are greatly under-utilized in current designs. Users can detect changes in size, color, shape, movement or texture [9]. Another possibility is a mix of the implemented features with the above suggested ideas. According to [9], we could think about a sequence of possible information representation (during a tele-guide session) following, as starting point, the *Visual Information Seeking Mantra*. This mantra is compound by seven tasks: overview, zoom, filter, details-on-demand, relate, history, extract. More information about *visual information seeking mantra* can be found in [9].

We could modify this sequence eliminating some tasks and redefining remaining ones. In particular the most important task for our purposes are the first three plus an optional one: overview, zoom, filter, details-on-demand.

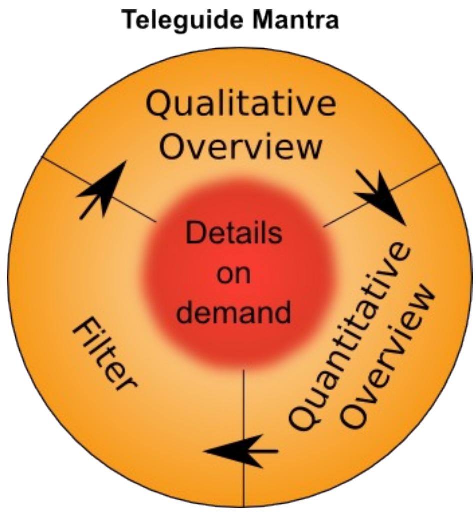


Image 11: Teleguide Mantra: sequence of the three steps plus the optional one (in the middle of the circle)

3.1 *Teleguide mantra*

Now, we give new definitions about these four task to better fit our purposes:

- **Qualitative overview**: a tele-operator should have a general perception of surrounding environment gathering generic information about distances between Morduc robot and other objects with few details. We think that a single frame with some coloured area should be enough to have an overview of the environment. We will call this step *qualitative overview.*
- **Quantitative overview**: for our purpose, we could change "zoom" word with *quantitative overview.* We mean that, inside the environment, there could be more than one POI[6]. In this step we could highlight only relevant POIs in a particular way. But what is a POI? A POI, in our test environment, is an object placed next to 3Morduc robot; in other words, a POI is a potential dangerous object because of its proximity to robot.
- **Filter**: when the operator has understood what surrounds the robot (through the two overview steps), superfluous details could be removed to clarify the view of the entire environment. For example, far objects should not be interesting for the tele-operator; so, after he has acquired an overview of the environment, far objects could not be superimposed by artificial graphics. At this point, only P.O.I.s are highlighted in a particular way. This choice let's user to view both distant safe areas and dangerous places (P.O.I.s) at the same time and real details of objects. Indeed, Shneiderman says that "*in 3-dimensional applications users must cope with understanding their position and orientation when viewing the objects*" [9].
- **Details-on-demand**: this is an optional step in the sense user could use it during tele-guide, during one of the other three steps without any limitation. This optional step let user to perform other orthogonal actions, e.g. focus (zoom, say) a particular area of the visualization

6 POI = point of interest

to see an enlargement of the scene (currently 3Morduc cameras have not focus features, so we should implement this feature in client), change the POV[7] of the scene.

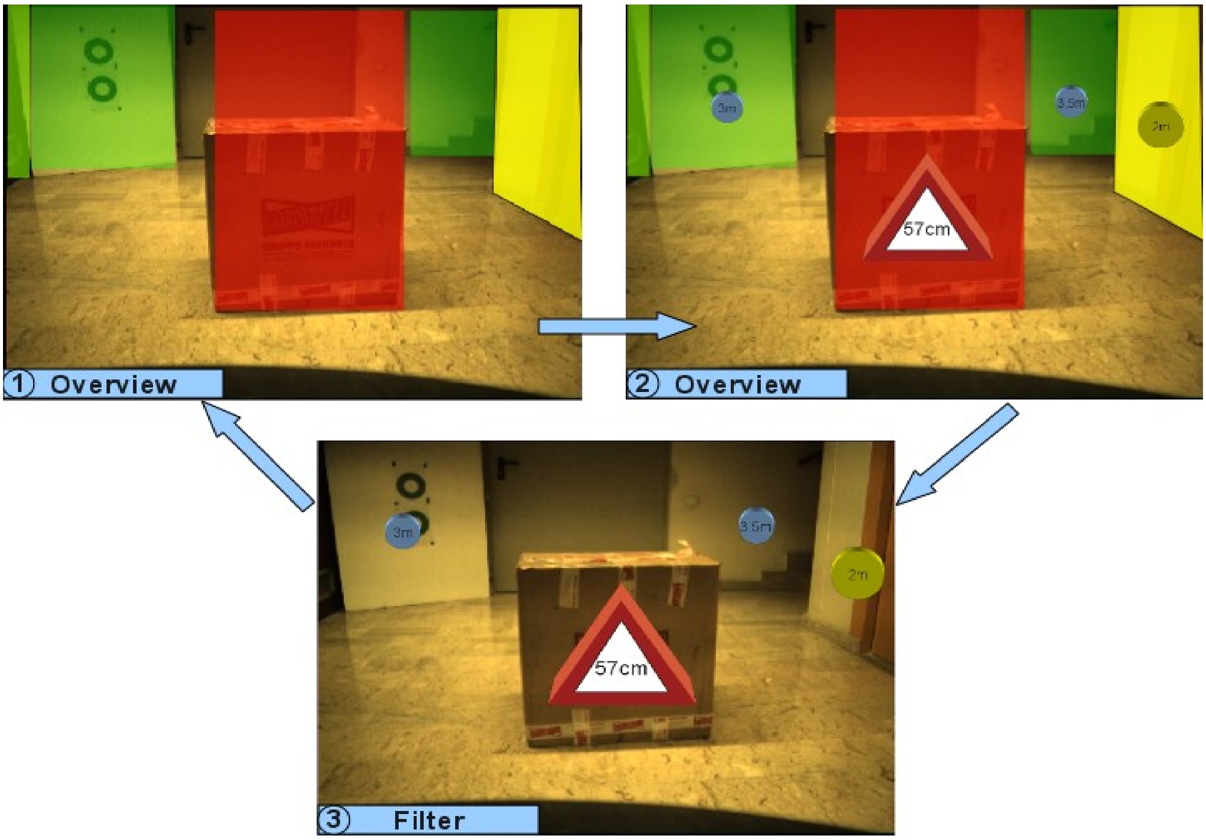


Image 12: Typical sequence using "Teleguide mantra": 1) Overview (qualitative overview), 2) Overview (quantitative overview), 3) Filter

In this way, after these three steps (plus, optionally, the fourth one), user should have the knowledge of:

- Entire environment;
- POIs (nearest objects = dangerous objects);
- Real details of close and distant objects.

The key point of the **Teleguide mantra** is to avoid user to be confused by the sequence of the three steps. We can, for example, skip some of them if user did not moved the robot or there is not any POI next to the robot (in this case Overview step could be skipped due to poor changes in the entire environment).

3.2 Using Teleguide mantra and client-side-prediction to decrease user perceived communication delay

We could use the *Teleguide mantra* to fill the delay provoked by our tele-guide system (see [1] for more details about system environment and delay reasons). In fact, we can use the delay to show the result of each task exploiting the inter-frame delay to communicate user more information about the environment. Maybe, these steps could not be enough because the mean delay could reach easily

7 Point Of View

more than few seconds between two following frames because of, for example, slow connections, long distances between robot and tele-operator (e.g. a space exploration mission). Shneiderman suggests in [9] the use of rapid, incremental, and reversible actions and the immediate display of feedback (less than 100 msec).

A solution to this delay problem could be the usage of the *Qualitative overview* step to predict the next frame information (see Image 10). If user requests robot to move, for example, we can try to predict the future distances and create a predictive overview of the entire environment. In this way, we suppose to fill the remaining delay presenting a false overview (not based on measured data, say), possibly similar to the real one, before the real *Quantitative overview* step result (which should be based on fresh laser data and new camera frames).

3.2.1 Client side prediction

This approach is called client-side prediction and it has been developed to solve similar delay problems in real time networked game such as Half-Life. In [17] a predictive modelling, which can be used to mask the effects of communication latency, is discussed by Bernier. The approach proposed by Bernier is applied to the multi-player server-based version of Half-life.

Typically a networked game like Half-Life[8] is based on a modified client/server infrastructure which is composed by an authoritative server that is responsible for running the main game logic

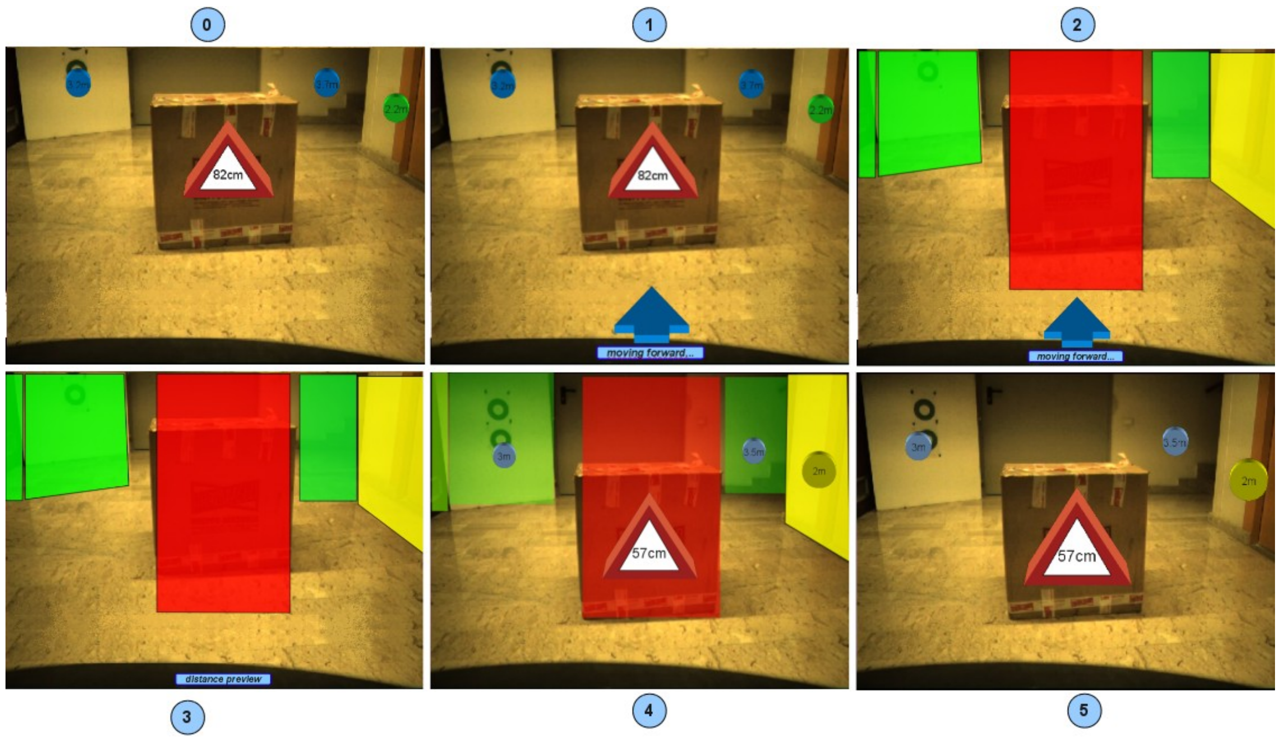


Image 13: An example of new data representation during tele-guide using Teleguide mantra and client-side-prediction. From image 0 to 5: previous filter step, user asks to go forward, robot moves forward and distance preview is shown, predictive overview waiting for fresh data, quantitative overview with fresh data, filter step

and one or more “dumb” clients connected to it. A peer-to-peer infrastructure is not usable because of different reasons: state of the game should be unique, a client should not provoke inconsistencies for other players, cheat avoidance. Bernier, in [17], describes a typical networked system for multi-

8 More information about Half-Life video game series can be found in [18]

player game: an authoritative server and one or more clients. "*Without client side prediction, the system works in the following way: clients sample user inputs and forward them to the server for execution; the server would execute the input commands, move around other objects and then send back to the client a list of objects to render*" [17]. Fiedler, in [19], says that "w*ith client side prediction, in order to remove the latency, the client runs more code than it previously did. It is no longer a dumb terminal sending inputs to the server and interpolating between state sent back. Instead it is able to predict the movement of your character locally and immediately in response to your input, running a subset of the game code for your player character on the client machine*".

Again Bernier, in [17], says that "*one method for ameliorating this problem is to perform the client's movement locally and just assume, temporarily, that the server will accept and acknowledge the client commands directly. There still is an authoritative server running the simulation just as noted above. Having an authoritative server means that even if the client simulates different results than the server, the server's results will eventually correct the client's incorrect simulation*".

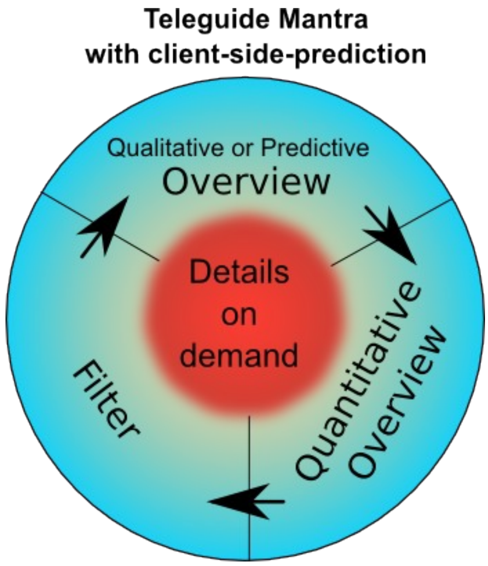


Image 14: Teleguide mantra sequence with client side prediction

We think that the proposed solution could fit our discussion because we have a similar system: an authoritative server placed inside 3Morduc and a client used by a tele-operator. In our test system, obviously, user movements requests do not modify 3Morduc inner state (which should describe objects in a digital world like a Half-Life authoritative server inner state) but it modifies the real placement of 3Morduc in the real world. 3Morduc client is executed in a powerful computer so it is possible to exploit its intelligence to predict something about 3Morduc future position after a movement request. As written above, using Teleguide mantra, we could substitute a qualitative overview step with a fictitious overview composed by the last valid camera frame as background and a predictive foreground based on computer graphic and other concepts discussed in previous paragraphs. So, we show user, for example, coloured stripes in predicted positions with predicted distances. In this way user could have an idea about future real distances as soon as possible during the time interval when the new camera frame (with fresh distance data) is produced by 3Morduc and sent to client.

When an up-to-date camera frame arrives to client, the temporary representation of the environment is substituted by new data and client and server are aligned again.

This approach should be useful because could help user to use delay period taking decision about future movements and let him to go back if an object is predicted to be too close.

So, using client side prediction, the Teleguide mantra needs to be modified, adding another parallel step to qualitative overview: predictive overview.

It is important to client-side-predict and show the prediction only in the first step (qualitative/predictive overview) of the Teleguide mantra because we are not able to know the correct values of object distances. We think that the written numbers (object distance, say) suggest users real data and not supposed ones.

4 Applying Teleguide mantra with client-side prediction to AR4Morduc client

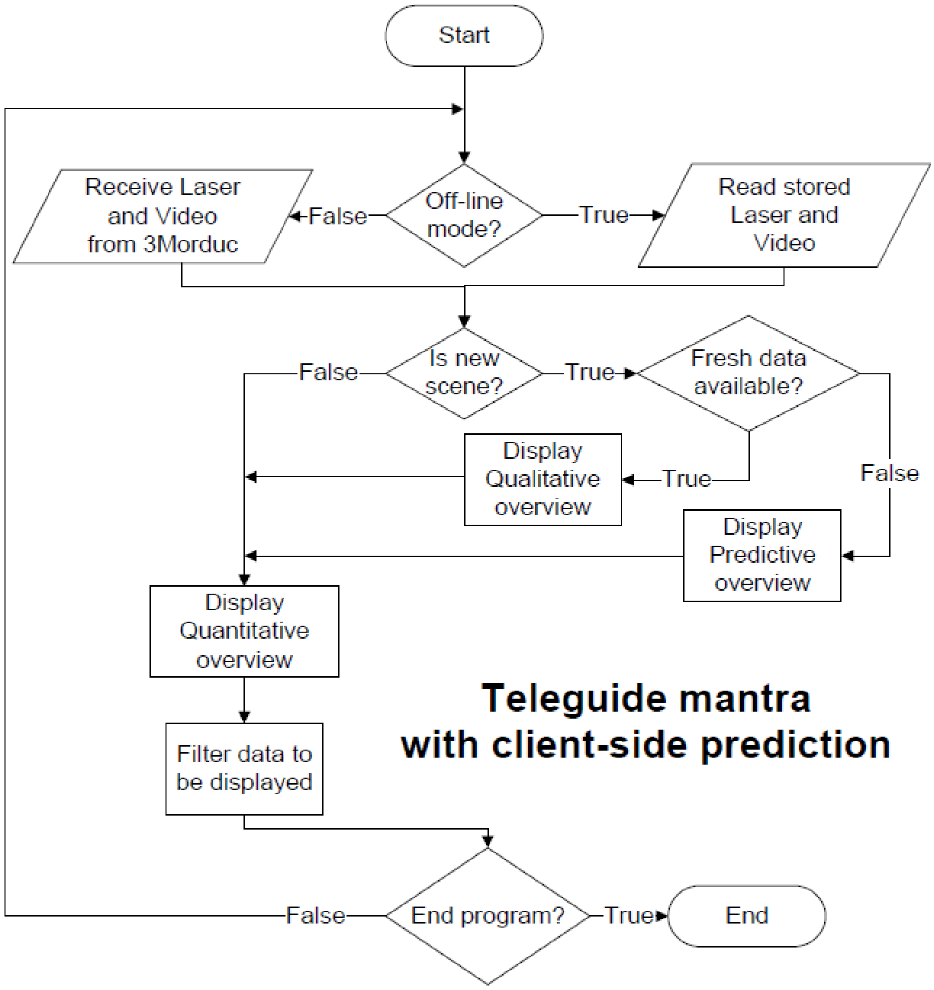
4.1 A Teleguide mantra with client-side prediction algorithm

Our goal is to present user a sort of VR superimposed on the last available real video frame until a valid video frame arrives to client. The concept of "validity" of the new video frame may be explained thinking about the sequence of the involved events:

1. client displays, using the *Filter* step of the *Teleguide mantra*, a valid video frame with some information superimposed (AR);
2. user asks for a movement;
3. user's request is sent to 3Morduc by the client;
4. 3Morduc moves;
5. 3Morduc send data (i.e. laser, video, odometer) to client
6. client receive fresh data to display using performing the *Teleguide mantra* overview step

Between step 3 and 6, some not up-to-date video frame could arrive to client because of communication latency. We could use the above suggested approach to avoid this misunderstanding.

The following flowchart explains the application of *Teleguide mantra with client-side prediction* discussed in the previous chapter.



Flowchart 1: Teleguide mantra with client-side prediction high-level flowchart

The concept of *validity* of a video image is connected to the time sequence of events that user expects: we should display only the most recent data and if it is not ready for displaying (not up-to-date video image and/or laser data, say), we provide a predictive representation of the environment instead of not up-to-date data.

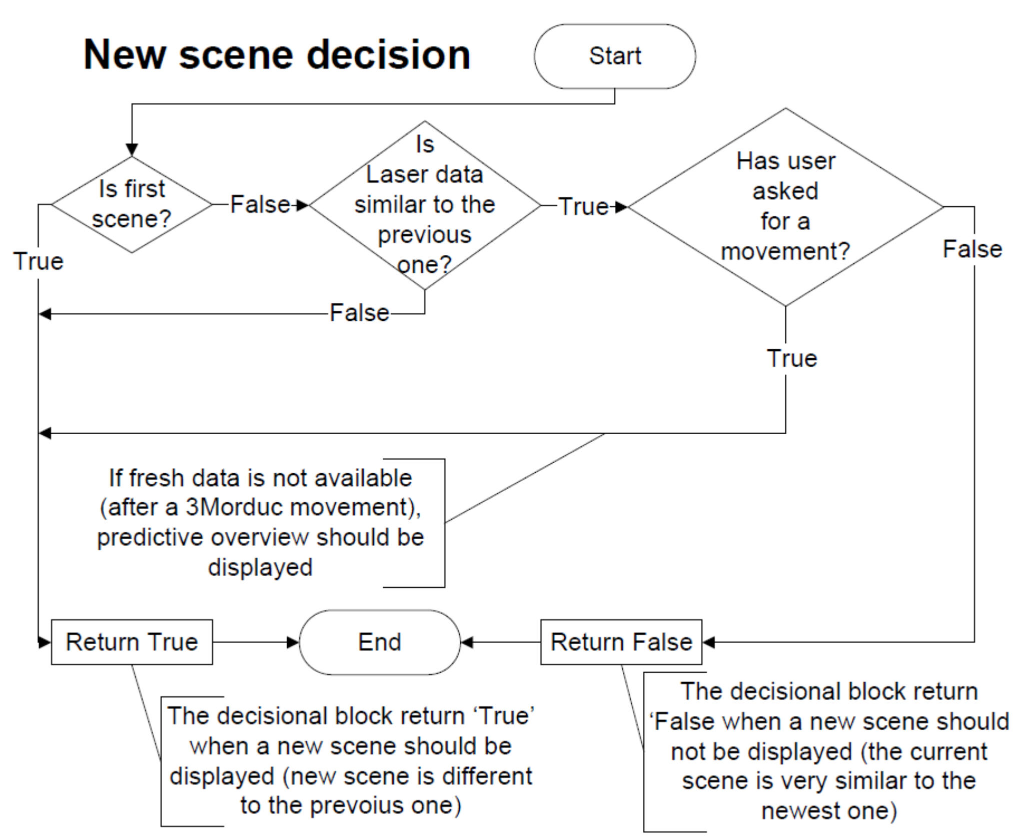
Two conditional blocks needs to be explained: *Is new scene* and *Fresh data available*.

4.1.1 "Is new scene" conditional block

Having a sequence of video frames and related laser data, it is important to not annoy user sight creating an effective and usable presentation of AR environment.

To do so, we could use to jump *Qualitative overview* and show directly *Quantitative overview*; this should be done only when a new scene is not arrived. We mean with "new scene" an arrived video frame and laser data which are different from the previous one[9].

9 Different approaches could be used to detect scene similarity.



Flowchart 2: New scene decision flowchart

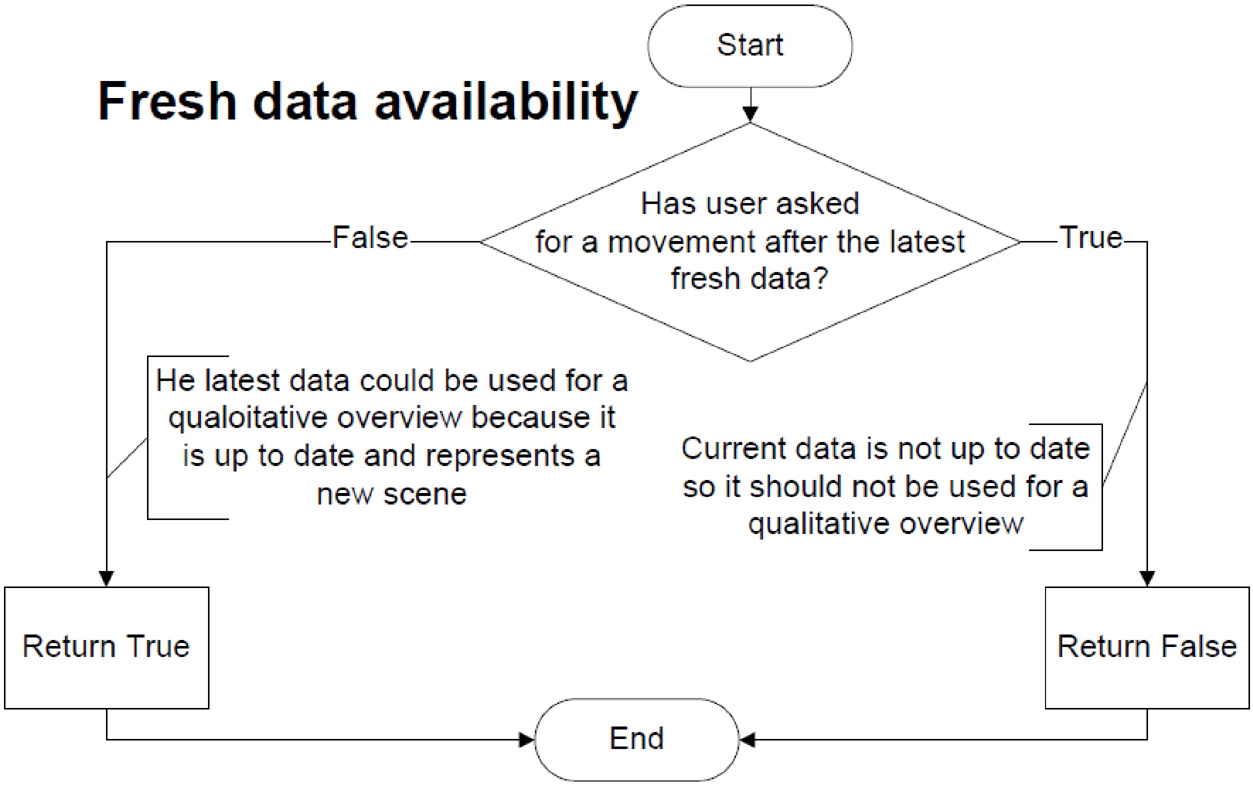
New scene decision should be based on key points:

- video frame similarity;
- laser data similarity;
- movement request by user.

As it is possible to read from the flowchart, if there is a poor similarity between video frames and/or laser data, or user requested for a movement, a new scene should be displayed; otherwise, the scene is similar to the previous one so a new overview step could be not useful.

4.1.2 "Fresh data available" conditional block

Supposing to receive a new video frame and related laser data, we must verify if they are "fresh". The concept of *freshness* is connected to the *request for a movement* and *data arrival relationship*: if user has not asked for a movement, arrived video frames and laser data are fresh; if user asked for a movement, arrived data could be not fresh even if they arrived after user's request.



Flowchart 3: Fresh data decision flowchart

The latter case is a tricky situation because we should consider user request time, new data arrival time, new data time-stamp and communication latencies. If we choose wrong settings, an erroneous display of not fresh data could happen: in this case not up-to-date data will be used to display one of Teleguide mantra steps.

4.2 Data representation using Teleguide mantra

We should organize distance representation helping user to understand in a short period of time a big quantity of information. In particular, according to MacDonald [21], "*key issues to consider are clarity, comprehensibility and how well the user will be able to pick out the desired information and understand its significance*". Other related factors are discernibility[10], conspicuity[11] and salience[12].

To enhance the above listed factors, we could use different approaches that could be mixed to create a more effective user experience. Tele-guide mantra offers many opportunities to experiment different way to display information.

4.2.1 Transparency

In AR applications the use of transparency is a common practice because it allows to add computer graphics layers over background images. Also MacDonald in [21] suggest to "*use color transparency to show overlays of related structures*". Our target is let user see the surrounding environment with distance information superimposed over a video image background; to do so we can use transparency in different ways according to the particular Teleguide mantra step.

10 How easy it is to distinguish an item from its background [21]
11 How obvious the item is relative to its neighbours [21]
12 How well the item *pops out* from the display as a whole [21]

In **Qualitative Overview** step we want to communicate user the distance information using bright coloured areas drawing the user attention on the entire environment and in particular on close objects (using a bright red); in this step it is important to let user have the perception of the entire distance environment structure. A low level of transparency should be appropriate because user attention will focused on coloured areas and not on background details.

In **Predictive Overview** step the client will predict and display a possible object placement using coloured areas and an old (not up-to-date, say) video image as background. In this case a correct level of transparency is a very low value because the displayed computer graphics environment is not real and background video image does not correspond to the current environment.

In **Quantitative Overview** step a middle level of transparency should be used because coloured areas are not the unique way to communicate distance information (solid icons or other object are superimposed over coloured areas). This step is a middle step between Qualitative/Predictive Overview and Filter steps so a kind of transition make user tele-guide experience more effective and comfortable.

In **Filter** step, transparency should not be used because the goal of this step is to allow user to see real details of the background video image. After the other steps, user has the knowledge of the entire environment distances and only important P.O.I. are highlighted by solid icons (or other foreground objects) communicating object distance.

Step	Level of foreground transparency (a value of zero means no transparency, a value of 100 means invisible object)[13]	Background image colour depth
Qualitative Overview	25	Colourful
Predictive Overview	15	Grey scale
Quantitative Overview	50	Colourful
Filter	100	Colourful

Table 2: Example values for level of foreground transparency and background image color depth for each Teleguide mantra step

4.2.2 Background video image

During Teleguide mantra steps some not up-to-date video images could be displayed (in particular during Predictive Overviews). There is a big difference between up-to-date and not up-to-date showed images because user should be aware about *freshness*. To make a distinction between these two kind of showed video images we can change the colour depth of not up-to-date video images converting them in a grey scale image before their visualization. In this way, foreground computer graphics coloured areas and object stand out over a not up-to-date image.

We think that a background video image should be always displayed even if it becomes not up-to-date: using Teleguide mantra, some environment predictions are showed and user should be aware about the most recent real environment condition, in this way tele-operator is able to understand better the meaning of the prediction.

13 Values should be tested, currently they have been written for example purposes only

5 References

1. S. Livatino, G. Muscato, D. De Tommaso, M. Macaluso, *Augmented Reality Stereoscopic Visualization for Intuitive Robot Teleguide*, 2010
2. B. L. Harrison, G. Kurtenbach, K. J. Vicente, *An Experimental Evaluation of Transparent User Interface Tools and Information Content*, 1995
3. D. Chang, L. Dooley, J. E. Tuovinen, *An Experimental Evaluation of Transparent User Interface Tools and Information Content*, 1995
4. J. Fultz, *Theory of Gestalt Psychology*, 1999 (broken link)
5. L. B. Ncube, A. W. Crispo, *Toward Meaningful Learning in A Global Age: How Gestalt Principles Can Facilitate Organization of Student Learning of Product Lifecycle Management Concepts*, 2007
6. M. Driscoll, *Psychology of learning for Instruction*, 1994
7. P. Baudisch, R. Rosenholtz, *Halo: a technique for visualizing off-screen objects*, 2003
8. Chandler, *Semiotics: The Basics, Routledge*, 2001.
9. B. Shneiderman, *The eyes have it: a task by data type taxonomy for information visualizations*, 1996
10. R. Williams, *The Non-Designer's Design Book*, 2007
11. M. A. Livingston and al., *Resolving Multiple Occluded Layers in Augmented Reality*, 2003
12. S. Livatino, G. Muscato, F. Privitera, *Stereo Viewing and Virtual Reality Technologies in Mobile Robot Teleguide*, 2009
13. S. Livatino, G. Muscato, S. Sessa, V. Neri, *DepthEnhanced Mobile Robot Teleguide based on Laser Images*
14. S. Livatino, G. Muscato, S. Sessa, C. Köffel, C. Arena, A. Pennisi, D. Di Mauro, E. Malkondu, *Mobile Robotic Teleguide Based on Video Images, Comparison Between Monoscopic and Stereoscopic Visualization*
15. V. M. L. Neri, *Mobile Robot tele-guide based on laser sensors*, 2008
16. Wikipedia, *Telepresence*, http://en.wikipedia.org/wiki/Telepresence, 2010
17. Y. W. Bernier, *Latency Compensating Methods in Client/Server In-game Protocol Design and Optimization*, 2001
18. Wikipedia, *Half-Life (series)*, http://en.wikipedia.org/wiki/Half-Life_(series), 2010
19. G. Fiedler, *What every programmer needs to know about game networking*, http://gafferongames.com/networking-for-game-programmers/what-every-programmer-needs-to-know-about-game-networking/, 2010
20. R. W. G. Hunt, *Measuring Colour*, 1998
21. L. W. MacDonald, *Using Color Effectively in Computer Graphics*, 1999
22. G. Murch, *Physiological Principles for the Effective Use of Color*, 1984
23. P. Thompson, K. May, R. Stone, *Chromostereopsis: A Multicomponent Depth Effect*, 1993
24. D. Chang, L. Dooley, J. E. Tuovinen, *Gestalt Theory in Visual Screen Design – A New Look at an Old Subject*, 2002
25. R. D. Zakia, *Perception and Imaging*, 1997