SGCS: Stereo Gaze Contingent Steering for Immersive Telepresence
Remi Cambuzat, Frederic Elisei, Gerard Bailly

To cite this version:
Remi Cambuzat, Frederic Elisei, Gerard Bailly. SGCS: Stereo Gaze Contingent Steering for Immersive Telepresence. Journee des Jeunes Chercheurs en Robotique (JJCR 2017), Nov 2017, Bidart, France. pp.1. hal-01677481

HAL Id: hal-01677481
https://hal.archives-ouvertes.fr/hal-01677481
Submitted on 8 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
1. Context

1.1 Research framework: Embodied Learning

Our goal is to “teach” a robot to interact autonomously in a face-to-face task with a human. Due to the complexity of the task, standard learning approach like learning by observation and kinesthetic demonstration are not efficient, notably for demonstrating social signals (such as gaze or head movements). Our approach consists in exploring the embodied learning paradigm, where a human pilot teaches the robot with his/her own moves [10]. Like a puppeteer, a pilot controls the robot remotely using an immersive social signal. In order to record an interaction with minimal bias, the platform should become “transparent” and the remote world represented in a “natural” way. Our research aims at providing the pilot with an active perception of the remote space, notably with a trustful and coherent perception of depth.

1.2 State of the Art

What we know about humans:

- Depth perception is based on several factors: Binocular disparity (stereovision), occlusion, parallax, convergence, known semantics of the objects, ... [1, 5, 9].
- Stereoscopic vision is useful before 15m (cannot differentiate from monovision after) [1].
- Vergence is useful in the peri-personal space (<2.0m) [14].

Current use of immersive teleoperation

- Search and rescue robot [9]
- Drone navigation [4]
- Stereoscopic vision is useful before 15m (cannot differentiate from monovision after) [1]
- Depth perception is based on several factors: Binocular disparity (stereovision), occlusion, parallax, convergence, known semantics of the objects, ... [1, 5, 9].

Gaze controlled methods

- Virtual gaze joystick: “Moving to the center” [17] [12]
- EyeSeeCam [11, 13]

Limitations of current immersive teleoperation devices

- Underestimation of depth in peripersonal space (<2m), overestimation after 2m [2].

Drawback: In those setups, the stereo rig is fixed. The pilot loses the vergence information/control, has reduced depth perception and experiences the accommodation-vergence conflict. For a human facing the robot, the robot gaze is less interpretable.

1.3 SGCS: Stereo Gaze Contingent Steering

Proposed approach:

Here we propose a new natural control method for a pair of stereoscopic robotic eyes with vergence abilities, called SGCS (Stereo Gaze Contingent Steering), running alongside the control of a robotic head. An evaluation of the control method has also been performed.

2. Technological platform

We use the Mical platform NINA from the CRISSP team at Gipsa. Specification :

- 1 lb 2.0 with enhanced face articulation [9].
- Cluster of 4 PC (3 Linux + 1 windows) using YARP (client-server robotic middleware).
- HTC Vive + SMI integration for eye-tracking
- IPD cameras equivalent to human IPD => reduced hyperstereopsis
- Communication with UDP/TCP through the YARP middleware
- The camera’s feeds are synchronised and displayed in the HMD (Head Mounted Display) as video texture.

3. Control methods

The control of the head and eye is done through a angular command for the six head encoders:

\[ \begin{align*}
\text{neck\_pitch} & = \text{neck\_pitch} \text{ encoders} \\
\text{neck\_roll} & = \text{neck\_roll} \text{ encoders} \\
\text{neck\_yaw} & = \text{neck\_yaw} \text{ encoders} \\
\text{eyes\_pitch} & = \text{eyes\_pitch} \text{ encoders} \\
\text{eyes\_roll} & = \text{eyes\_roll} \text{ encoders} \\
\text{eyes\_yaw} & = \text{eyes\_yaw} \text{ encoders}
\end{align*} \]

1. Head control: The head angles are driven by the HMD orientation value (standard approach).
2. Eye control: The gaze information returned (in pixel) is converted in UV coordinates, relative to the displayed video texture referential.
3. Inverse model: Using a transfer matrix \( UV_{\text{camera}} \) – that links the ROI placement on the rectified stereo camera images with the angular values that would align the ROI with the cameras’ optical axis –, this inverse model carries over the pilot’s gaze direction to the robot’s eyes movements. This inverse linear model has (surprisingly) a precision of 0.5° on the three angles.
4. Forward display: We then move the center of the video texture to a new UV coordinates pair, calculated by the Forward model from the eyes encoders angular values. This moves the video texture in the virtual world to a coherent position for the robot and cues on the pilot side.

4.1 Setup & protocol

Setup: 7 target at various distance (25 to 100cm)
Protocol: For the reference condition (Ideal target angles determined semi-automatically) and the pilot, every target has been seen 8 times. On 4 passes (left->right, front->back, right->left, back->front) repeated twice times.
Subjects: 16 subjects (3 women, 13 men), aged between 22-56 yo. No prior experience of virtual reality before for most of them (13 inexperienced VR).

4.2 Results

Discussion: Our SGCS control method is able to move the robotic eye in coherence with the orientation of the human eye (the cameras’ optical axes are aligned with the human gaze). The cameras are looking where the human is looking with respect of tilt, azimuth, and vergence.

Future works:

- Hypothesis: Control of vergence improves perception and evaluation of depth in the near and medium field while maintaining oculomotor cues and reducing the accommodation-vergence conflict.
- Improve the reactivity of the control method: detection of fixation and saccade.

References